PREDICTIVE MODELING OF HIGH-CURRENT OUTPUT RESISTANCE AND THERMAL EFFECTS IN BIPOLAR JUNCTION TRANSISTORS

By

SANG-GUG LEE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1992

UNIVERSITY OF FLORIDA LIBRARIES

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my adviser, Professor Robert M. Fox, for his guidance, encouragement, and support throughout the course of this work. I would also like to thank the other members of my committee, Professors Jerry G. Fossum, Dorothea E. Burk, William R. Eisenstadt, and Kevin S. Jones, for their interest in this research and patience in reading through the manuscript.

I am grateful to Delco, Harris Semiconductor, Motorola, and National Semiconductor for providing test devices and technical advices that made this work possible. The financial support of the Semiconductor Research Cooperation is gratefully acknowledged.

I would like to thank Bill Davis of Motorola, whose work in thermal feedback [Da88] encouraged us to pursue the research on thermal modeling, and David Zweidinger, who has implemented the thermal model into SPICE code.

Thanks are extended to many of my colleagues and friends, in particular, Jin-Young Choi, Yeong-Seuk Kim, Joo-Hyun Jin, Han-Jin Cho, Ghy-Boong Hong, Hae-Seok Cho, Yungseon Eo, Tae-Hoon Kim, Dong-Wook Suh, Hee-Myong Park, Keith R.

Green, D. Scott Langford, Jim Jaffee, Deodatta R. Apte, and Minchang Liang for helpful discussions and encouragements.

My deepest gratitude goes to my parents and parents-in-low for their endless love. Last but not the least, I thank my wife Myong-Suk my daughter Si-Eun whose endless love and encouragements have sustained me throughout the years of my graduate study.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS ii

ABSTRACT vi

CHAPTERS

1 INTRODUCTION 1

2 THE EFFECTS OF CARRIER-VELOCITY SATURATION ON HIGH-CURRENT BJT OUTPUT RESISTANCE 10

2.1 Introduction 10
2.2 Analysis 11
2.2.1 Case 1 14
2.2.2 Case 2 21
2.2.3 Case 3 22
2.3 Numerical simulations 25
2.4 Conclusions 29

3 PREDICTIVE MODELING OF THERMAL EFFECTS IN BJTS ... 32

3.1 Introduction 32
3.2 Model for thermal spreading resistance 33
3.3 Small-signal model 44
3.4 AC modeling 48
3.5 Conclusion 50

4 THERMAL-PARAMETER EXTRACTION FOR BIPOLAR CIRCUIT MODELING..... 54

4.1 Introduction 54
4.2 Thermal impedance extraction methods 55
4.2.1 Using output resistance as a temperature sensitive parameter 56
4.2.2 Using reverse transconductance as a temperature sensitive parameter .. 60
4.3 Extraction results 63
4.4 Conclusion 68

5 THERMAL EFFECTS IN BJT INTEGRATED CIRCUITS 69

5.1 Introduction 69
5.2 Modified SPICE BJT model 71
5.3 Thermal effects in integrated circuits 82
5.3.1 DC thermal effects in ICs 83
5.3.2 AC thermal effects in ICs 99
5.4 Summary 121

6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK 124

REFERENCES 127

BIOGRAPHICAL SKETCH 131

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

PREDICTIVE MODELING OF HIGH-CURRENT OUTPUT RESISTANCE AND THERMAL EFFECTS IN BIPOLAR JUNCTION TRANSISTORS

BY

SANG-GUG LEE

August 1992

Chairman: Dr. R. M. Fox
Major Department: Electrical Engineering

This dissertation concerns predictive modeling of high-current output resistance and thermal effects in bipolar junction transistor (BJT). Modern BJTs tend to operate with saturated carrier velocity in the collector space-charge region (SCR). A physical analysis of the effects of velocity saturation on the output resistance of BJTs is presented, especially for high current levels. Physical analyses show that when the collector current density approaches a critical value, the Early voltage can increase significantly due to carrier-velocity saturation as the resulting base push-out partially offsets base-width modulation. This effect is also demonstrated in simulations with PISCES and MMSPICE. The predicted peak in Early voltage cannot be observed in low-frequency measurements due to the thermal effects. The

effects of self-heating on BJT behavior are demonstrated through measurement and simulation. Most affected are the small-signal parameters Y_{22} and Y_{12}. A frequency-domain solution to the heat-flow equation is presented which applies to any rectangular emitter geometry. The model, although simple enough for CAD, predicts thermal spreading impedance with good accuracy for a wide range of frequencies. A practical method is presented for extracting thermal spreading resistance of BJTs. The method uses the output resistance or reverse transconductance as the temperature-sensitive parameter. Measurement can be made in the time or frequency domain. General insights into local-heating effects on the large- and small-signal behavior of integrated circuits are provided, and the circuit conditions under which self-heating causes significant errors in simulation are identified. A thermally improved BJT model is used as a verification tool in studying the implication of thermal effects on integrated circuits.

CHAPTER 1
INTRODUCTION

Analog circuits such as operational amplifiers typically employ active loads to achieve high gains. The voltage gain and frequency response of such circuits thus depends on the transistors' small-signal output resistance. In most contemporary bipolar junction transistors, high electric fields occur in the collector junction space-charge region due to the high epi-layer doping or high reverse bias voltage. These high fields tend to saturate the free-electron drift velocity, and thereby affect output resistance.

In Chapter 2 of this dissertation, a study is done of the effects of carrier-velocity saturation, especially for high current levels, on the output resistance characteristics of BJTs. Physical and numerical analyses show that when the collector current density approaches a critical value, the Early voltage can increase significantly due to carrier-velocity saturation as the resulting base push-out partially offsets base-width modulation. This effect is demonstrated in simulations with PISCES and MMSPICE. Measurements suggest that thermal feedback tends to mask this effect at low frequencies. The discrepancies between analysis and low-

frequency measurements led us to study the thermal effects in bipolar junction transistors.

Existing electronic circuit simulation programs, both empirical (Gummel-Poon, SPICE 2 and 3) and physics-based (PISCES, MEDUSA, BIPOLE, MMSPICE), generally model the integrated circuit as if all the devices within the circuit are at a constant temperature. In fact, the power dissipated by the elements of the integrated circuit can raise the instantaneous temperature of the element itself and other elements above ambient temperature. Increases in device temperature can lead to significant increases in terminal currents and/or decreases in base-emitter voltage depending on the bias conditions. The thermally induced variations can also lead to significant changes in y-parameters, especially Y_{12} and Y_{22}.

There has been a considerable amount of work done on the modeling of thermal feedback in BJTs [Sp58, St59, Mu64, Jo70, Me77, Da88, De89], yet these effects are neglected in most existing device and circuit simulators. One probable reason for this is the empirical approach which is popular in BJT modeling. For example, the Early voltage, commonly used to model output resistance, actually drops at high currents, yet a constant value is still commonly used. A second reason is that thermal effects are considered to be a problem unique to power transistors.

The temperature rise in a device is proportional to the thermal resistance, defined as the total temperature rise caused by a unit power step. A packaged transistor can be characterized by two thermal impedance components, viz., the thermal spreading impedance controlling the fast temperature variation and the package thermal impedance controlling the slow process. The package thermal impedance is dependent on the thermal properties of the packaging material and cause coupling of all the devices on the chip. The thermal spreading impedance is concerned with the geometry of the transistor. In the past, the contribution of the thermal spreading resistance to the total thermal impedance has been negligible. However, circuit and device integration levels continue to increase with time. The resulting smaller device size leads to higher thermal spreading impedance so that it tends to dominate the total thermal impedance. At the same time, the demand for high performance has led to higher power densities in these devices. The result of these trends is towards larger temperature rises in device. Therefore, in modern high-performance circuits, the device thermal resistances and power densities are sufficient to result in substantial temperature rises within a circuit. Also, new isolation technologies (e.g., trench isolation, dielectric isolation, and SOI) tend to make thermal effects more

The temperature rise in a device is proportional to the thermal resistance, defined as the total temperature rise caused by a unit power step. A packaged transistor can be characterized by two thermal impedance components, viz., the thermal spreading impedance controlling the fast temperature variation and the package thermal impedance controlling the slow process. The package thermal impedance is dependent on the thermal properties of the packaging material and cause coupling of all the devices on the chip. The thermal spreading impedance is concerned with the geometry of the transistor. In the past, the contribution of the thermal spreading resistance to the total thermal impedance has been negligible. However, circuit and device integration levels continue to increase with time. The resulting smaller device size leads to higher thermal spreading impedance so that it tends to dominate the total thermal impedance. At the same time, the demand for high performance has led to higher power densities in these devices. The result of these trends is towards larger temperature rises in device. Therefore, in modern high-performance circuits, the device thermal resistances and power densities are sufficient to result in substantial temperature rises within a circuit. Also, new isolation technologies (e.g., trench isolation, dielectric isolation, and SOI) tend to make thermal effects more

significant due to the poor thermal conductivity of silicon dioxide.

A well-known attribute of the flow of heat from a small source into large media is its effective confinement to a region less than a few times the lateral dimension of the source [Ho58]. For transistors dissipating large power, heating within the elements of the integrated circuit can cause considerable chip temperature gradients. In this case, the power dissipation not only influences the performance of the device itself but leads to thermal coupling between devices. This type of thermal coupling is well known and in most cases they can be avoided by using symmetrical layout, placing in close proximity, or arranging critical pairs of transistors as cross-connected quads. For modest powers, even adjacent devices are thermally decoupled from each other [Fo91a] so that localized heating of the BJT is intrinsic to the device itself and fundamental. This dissertation focuses on this intrinsic self-heating effect which cannot be avoided through changing the layout of the IC.

Joy and Schlig [Jo70] produced an excellent time-domain model for the thermal response. In their time-domain derivation, Joy and Schlig [Jo70] integrated a point-source heat-flow equation over the volume of the collector space-charge region (SCR) and its image. The result is a closed-form expression for the impulse response at any point in the

semiconductor. The time integral of this expression gives the response to a unit step in the collector power. The integral is complicated and does not have a closed-form solution. The step response at $t = \infty$ gives the thermal resistance R_{th}. Their model is too CPU-intensive for use in circuit simulation. Furthermore, thermal feedback tends to cause more significant errors in the small-signal behavior of a circuit than in the transient response. A simple, yet scalable and predictive frequency-domain model of thermal spreading impedance was needed for small-signal thermal modeling.

Recently several workers [Vo89, Ye91] have demonstrated ways to model the thermal feedback in circuit simulation assuming thermal model parameters are available. While the thermal spreading resistance is the dominant part of the total thermal impedance in scaled-down devices, characterizing these devices is very difficult. Based on Joy and Schlig's model, thermal time constants of the scaled devices are on the order of nanoseconds. The standard electrical method for thermal impedance measurement is the "emitter-only switching" method, and it is adapted for characterizing power transistors [Oe76]. In this technique, a large heating current is first applied to establish thermal equilibrium. This current is then turned off and the temperature is inferred from the resulting transient response of V_{BE}, measured at a small fixed emitter current. The initial

response is controlled by both electrical and thermal spreading-impedance effects. This makes it difficult to determine the original temperature, and prevents accurate extraction of the thermal spreading resistance R_{th}. Therefore, there was a need for more reliable characterization method for thermal parameters.

It was shown [De89] that neglecting intrinsic local heating effects caused errors up to 4% in the predicted delay of a simple BJT logic gate. Local heating can severely degrade the performance of certain types of analog integrated circuits. The types of circuits that are subject to thermal effects are those circuits which require high precision such as D/A converters, instrumentation amplifiers, analog multipliers, logarithmic amplifiers, current mirrors, and precision voltage references, and those types of circuits which experience large amounts of power dissipation such as voltage regulators, audio and servo amplifiers. Local heating can also cause errors in op-amp practice as it can lower the low-frequency open-loop gain, cause distortion, change the phase response as well as the offset voltage and common-mode-rejection of differential amplifiers. Therefore, it is important for designers to have good insight into the circuit implications of local-heating effects, as even well-executed designs can fall prey to these effects [To90]. Knowledge of the implications of thermal effects in integrated circuits

not only provides the design principles to reduce thermal effects in IC design but can be used to exploit them for better IC design.

This dissertation, then, is concerned with predictive modeling of thermal effects in bipolar junction transistors and the circuit implications of thermal effects. This work will provide a good insight on the high current output-resistance characteristics and a general understanding and prediction of thermal effects in BJT circuits. The main contributions made in this work are as follows:

(1) predictive modeling of the effects of carrier-velocity saturation, especially for high current levels, on the output-resistance characteristics of BJTs;

(2) a simple and yet predictive frequency-domain thermal impedance model;

(3) measurement techniques for characterizing thermal impedance;

(4) general insights into local-heating effects on the DC and small-signal performance of integrated circuits and identifcation of the circuit conditions under which self-heating causes significant errors in simulation.

Chapter 2 presents a physical analysis of the effects of velocity saturation on the output resistance of BJTs. Analytical expressions for normalized BJT output resistance

are derived for an idealized structure. MMSPICE and PISCES-II simulations are also provided.

In Chapter 3, the effects of self-heating on BJTs are demonstrated through measurement and simulation. This chapter will show that local-heating causes much larger errors in BJT small-signal parameters than large-signal parameters. A thermal impedance model is derived that predicts thermal corrections to BJT small-signal parameters based on a solution of the heat-flow equation in the frequency domain. The model is applicable in principle to any emitter geometry. Polynomial approximations provide simplified expressions for rectangular emitters. These expressions are simple enough for compact modeling in SPICE.

In Chapter 4, practical methods are presented for extracting the thermal spreading resistance of BJTs separately from package thermal impedance. The methods use the output resistance or the reverse transconductance as the temperature-sensitive electrical parameter. Measurement techniques in both time and frequency domain are introduced.

The thermal impedance model described in Chapter 3 has been implemented in SPICE DC and AC analysis by D. Zweidinger [Zw92]. In Chapter 5, the improved SPICE BJT model [Fo91c] is introduced which includes a brief description of how the thermal models are incorporated in the DC and AC routines, and discusses the three levels of thermal modeling available

to the user. In addition, a series of simulations are compared to measurements to demonstrate the improved accuracy afforded by this thermally enhanced BJT model. DC thermal effects are considered in two aspects, magnification of the increments in collector current due to the mismatch in collector-base voltage and collector-current deviation from near-exact exponential dependence on base-emitter voltage. On the small-signal side, the relations between the thermal effects on the input and output resistances and the series emitter (R_E), base (R_B), and collector (R_C) resistances are discussed. The implication of discrepancies in the input and output resistance to other small-signal aspects are also discussed. Simulation results of a few circuit configurations, based on modified and ordinary SPICE, are provided for verification and demonstration purposes. Design strategies to avoid significant thermal problems are also discussed in this chapter.

In Chapter 6, the main accomplishments of the dissertation are summarized, and suggestions of areas for future work are made.

CHAPTER 2
THE EFFECTS OF CARRIER-VELOCITY SATURATION ON HIGH-CURRENT BJT OUTPUT RESISTANCE

2.1. Introduction

Analog circuits such as operational amplifiers typically employ active loads to achieve high gains. The voltage gain and frequency response of such circuits thus depend on the transistors' small-signal output resistance. In most contemporary bipolar junction transistors, high electric fields occur in the collector junction space-charge region due to the high epi-layer doping or high reverse bias voltage. These high fields tend to saturate the free-electron drift velocity, and thereby affect output resistance.

This chapter presents a study of the effects of carrier-velocity saturation, especially for high current levels, on the output resistance characteristics of BJTs. Analytical expressions for normalized BJT output resistance are derived for an idealized structure as an aid to qualitative understanding. MMSPICE [Je89b] and PISCES-II [Pi85] simulations are also provided.

2.2. Analysis

In a recent letter, Roulston [Ro90] presented an improved model for the Early voltage, V_A, in NPN transistors with very narrow base-widths (less than 0.1μm). Roulston concluded that, due to carrier-velocity saturation, a substantial improvement in Early voltage is expected in narrow-base transistors compared to estimates based only on base doping level and width. Liou's [Li90] correction to this work led to the following:

$$V_A = V_{A0} \left[1 + D_n F_c / v_s W_{B0}\right] \tag{2.1}$$

where $V_{A0} = W_{B0}/(dW_{B0}/dV_{CE})$ is the conventional expression for Early voltage (neglecting velocity saturation), D_n is the electron diffusion constant, F_c is an empirical factor to account for built-in fields and bandgap narrowing in the base region, v_s is the electron saturation velocity, and W_{B0} is the quasi-neutral base width. For a silicon bipolar transistor with a base width of 0.04 μm, D_n = 17 $cm^2/V{\cdot}s$, and F_c = 1, this model predicts $V_A = 1.4\ V_{A0}$.

With carrier-velocity saturation, the carrier concentration in the space-charge region (SCR) is proportional to the collector current, and the positions of the SCR boundaries become functions of the carrier concentration. Roulston's model accounts only for a reduction in the concentration gradient due to the nonzero carrier

concentration at the SCR boundary. From (2.1), this leads to significant increase in V_A only for very narrow base widths. As is well known, at the level of collector current where the SCR boundaries become strong functions of J_C, base-width modulation due to increments in V_{CB} tends to be offset by base push-out (Kirk effect) [Ki62] as the resulting increase in n_c (and J_C) redistributes the electric fields in the SCR. This offsetting mechanism, which was neglected in [Ro90], can lead to a much more significant increase in small-signal output resistance at high currents, even for ordinary base-widths.

Following is a derivation of this increase in output resistance as represented by the Early voltage. The Early voltage [Ge78] can be considered as the output resistance normalized by the collector current: $V_A = r_o I_C - V_{CE}$ where r_o is the output resistance and I_C is the collector current. To simplify the mathematics, base and buried-layer dopings are assumed constant, and the buried-layer doping is assumed much greater than the base doping, which is much greater than the epi-doping. The base terminal is assumed to be driven by a constant voltage source (Fig. 2.1) and the emitter junction SCR width is assumed to be independent of the collector current. The base-collector bias voltage and the resulting electric field are assumed to be high enough to saturate the electron velocity even near the SCR edge. By this assumption we are excluding the possibility of ohmic quasi-saturation

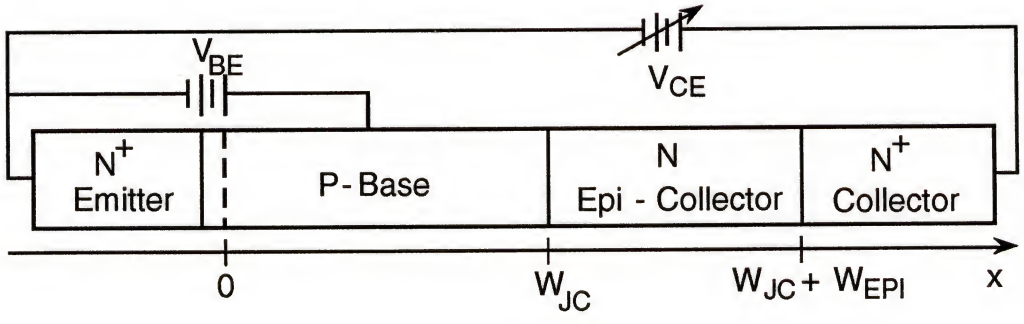


Fig. 2.1. One-dimensional NPN transistor structure.

even though in practice the transistor might enter that operating region at low V_{CB}. Ohmic quasi-saturation is excluded from this analysis since it can be treated without considering velocity saturation.

Three operating conditions are considered: (Case 1) forward-active with part of the epitaxial layer quasi-neutral, (Case 2) forward-active with the epi-layer entirely space-charged, and (Case 3) non-ohmic quasi-saturation [Je89a].

2.2.1 Case 1

Figure 2.2a shows the electron concentration and electric field distribution for the case of forward-active operation with part of the epi-layer quasi-neutral (QN). In the SCR, the current flows predominantly by drift. Assuming velocity saturation, the collector current density J_C can be expressed as

$$J_C = q n_c v_s \tag{2.2}$$

where n_c is the electron concentration in the space-charge region required to support the current. Equation (2.2) and the equality of charges on both sides of the collector junction lead to the relation

$$x_B(q N_A + J_C/v_s) = x_C(q N_{EPI} - J_C/v_s) \tag{2.3}$$

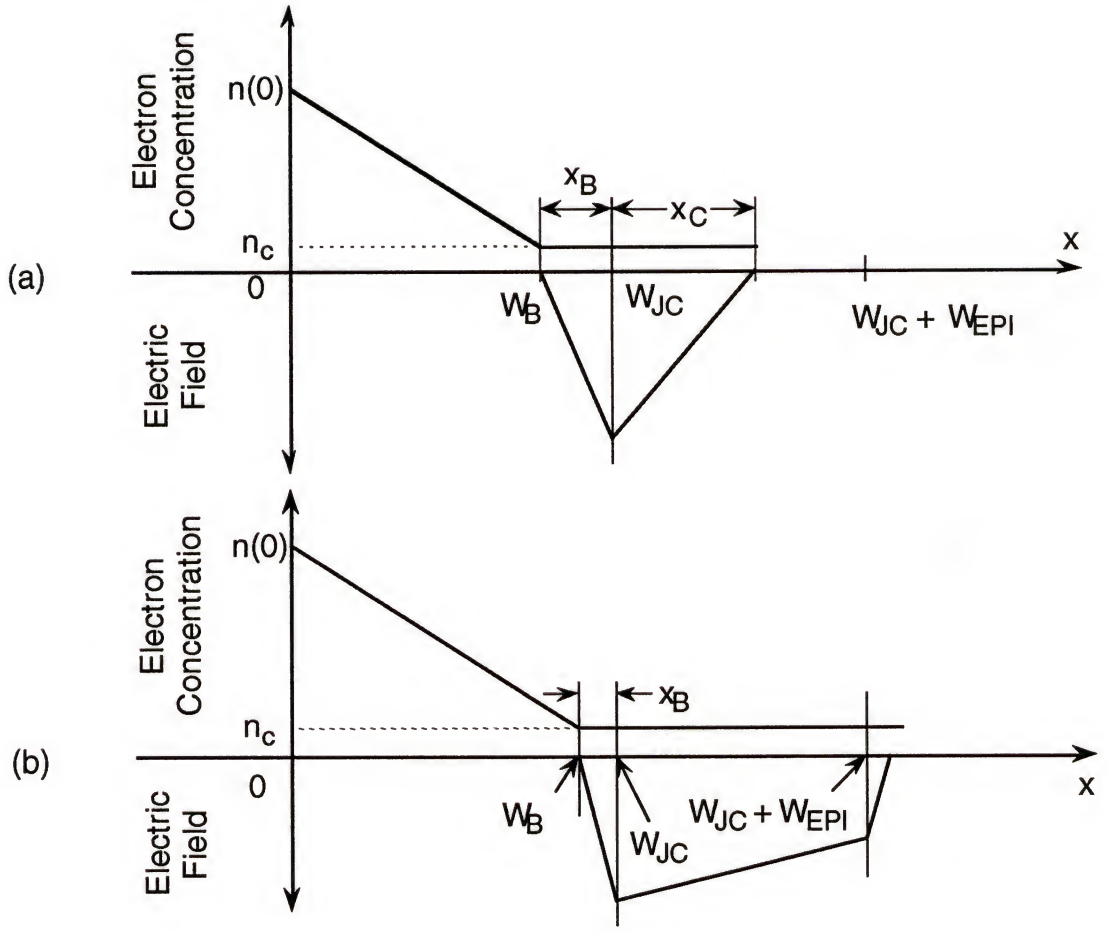


Fig. 2.2. Typical electron concentration and electric field distribution for (a) forward-active operation with part of the epitaxial layer quasi-neutral. (b) forward active operation with the epi-layer entirely space-charged, $J_C < J_{EPI}$. (c) forward active operation with the epi-layer entirely space-charged, $J_C > J_{EPI}$. (d) non-ohmic quasi-saturation operation.

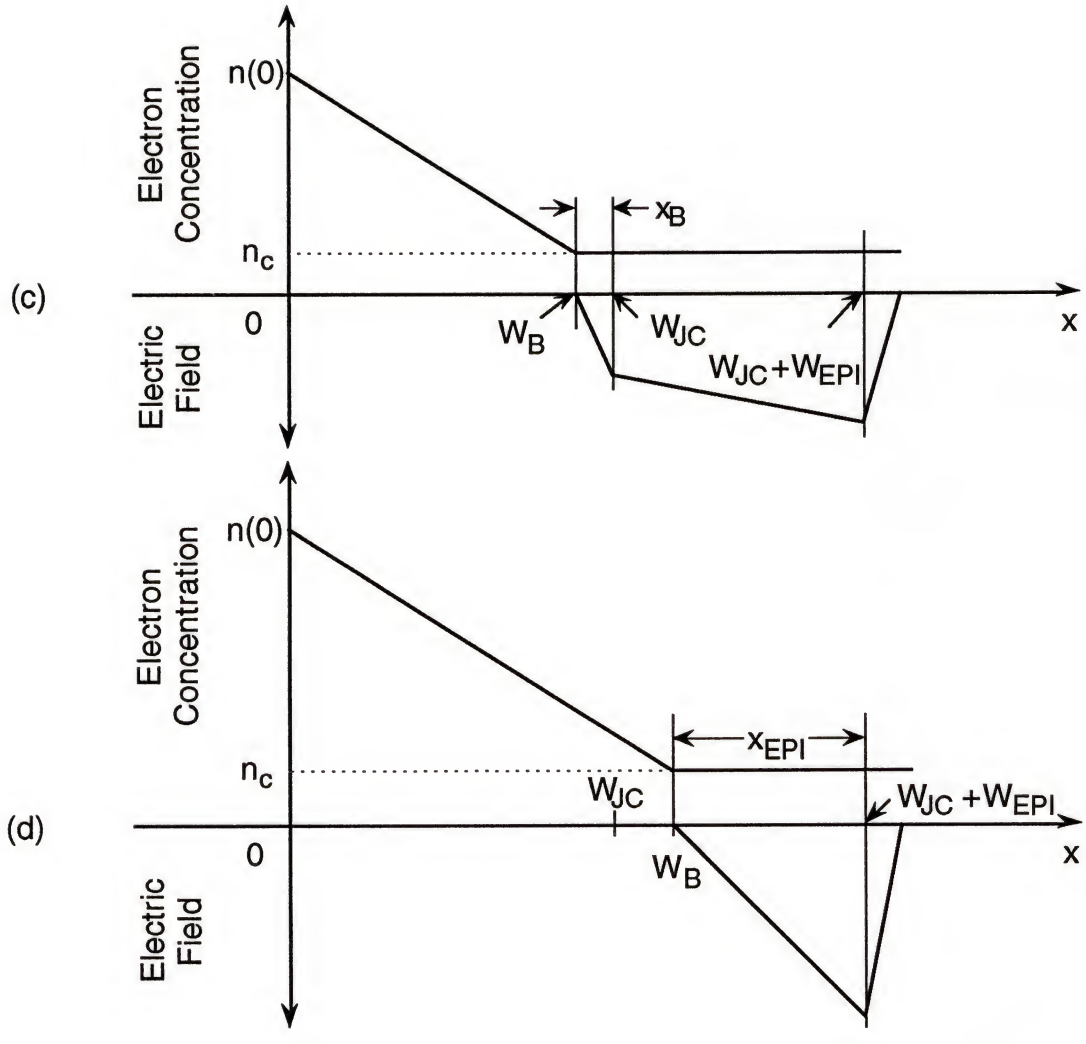


Fig. 2.2 continued

where x_B is the penetration of the SCR into the base, x_C is the penetration of the SCR into the epi-layer, N_{EPI} is the epi-layer doping concentration, and N_A is the base doping concentration. Poisson's equation can be solved in the collector junction SCR to yield the following expression:

$$V_{CB} + \phi_C = \frac{1}{2\varepsilon_s}\left[\left(qN_A + \frac{J_C}{v_s}\right)x_B^2 + \left(qN_{EPI} - \frac{J_C}{v_s}\right)x_C^2\right] \qquad (2.4)$$

where V_{CB} is the collector junction reverse bias, ϕ_C is the collector junction built-in potential, and ε_s is the semiconductor permittivity. In (2.4) the potential drop across the quasi-neutral portion of the epi-layer is neglected for simplicity. The effect of this approximation on output resistance will be discussed later. From (2.3) and (2.4), x_B can be written in terms of V_{CB} and J_C as:

$$x_B = \sqrt{\frac{2\varepsilon_s N_{EPI}\,(V_{CB}+\phi_C)}{qN_A\,(N_{EPI}+N_A)}}\sqrt{\frac{1-J_C/J_{EPI}}{1+J_C/J_A}} = x_{B0}\sqrt{\frac{1-J_C/J_{EPI}}{1+J_C/J_A}} \qquad (2.5)$$

where $$x_{B0} = \sqrt{\frac{2\varepsilon_s N_{EPI}\,(V_{CB}+\phi_C)}{qN_A\,(N_{EPI}+N_A)}}\,, \qquad (2.6)$$

$J_{EPI} = qN_{EPI}v_s$, and $J_A = qN_Av_s$. x_{B0} is the SCR penetration into the base which would result from neglecting velocity saturation (i.e., using the depletion approximation).

The quasi-neutral base-width (W_B) with carrier-velocity saturation can be expressed by

$$W_B = W_{JC} - x_B \tag{2.7}$$

while the base width neglecting velocity saturation (W_{B0}) is expressed by

$$W_{B0} = W_{JC} - x_{B0}. \tag{2.8}$$

Figure 2.3 shows a plot of x_B/x_{B0} as a function of collector current density J_C. As can be seen from Fig. 2.3, as J_C approaches J_{EPI}, x_B decreases rapidly. Therefore, dW_B/dV_{CE} differs significantly from the value dW_{B0}/dV_{CE} found in Roulston [Eqn. 2.1].

J_C can be expressed as

$$J_C = \kappa q D_n [n(0) - n_c]/W_B \tag{2.9}$$

where κ is an empirical factor that allows for drift current and $n(0)$ is the electron concentration at the edge of emitter-base SCR in the base region (Fig. 2.2a). Using (2.2), (2.9) and the definition of the Early voltage, V_A can be expressed as

$$V_A = -[W_B + \kappa D_n/v_s](dV_{CE}/dW_B). \tag{2.10}$$

Therefore, from (2.5), (2.7), and (2.10), we obtain

$$V_A \cong \frac{W_B(V_{CB} + \phi_C)}{0.5\ x_B}\left[1 + \frac{\kappa D_n}{v_s W_B} + \frac{\varepsilon_s v_s (V_{CB} + \phi_C) J_C}{x_B W_B (J_C + J_A)^2}\right]. \tag{2.11}$$

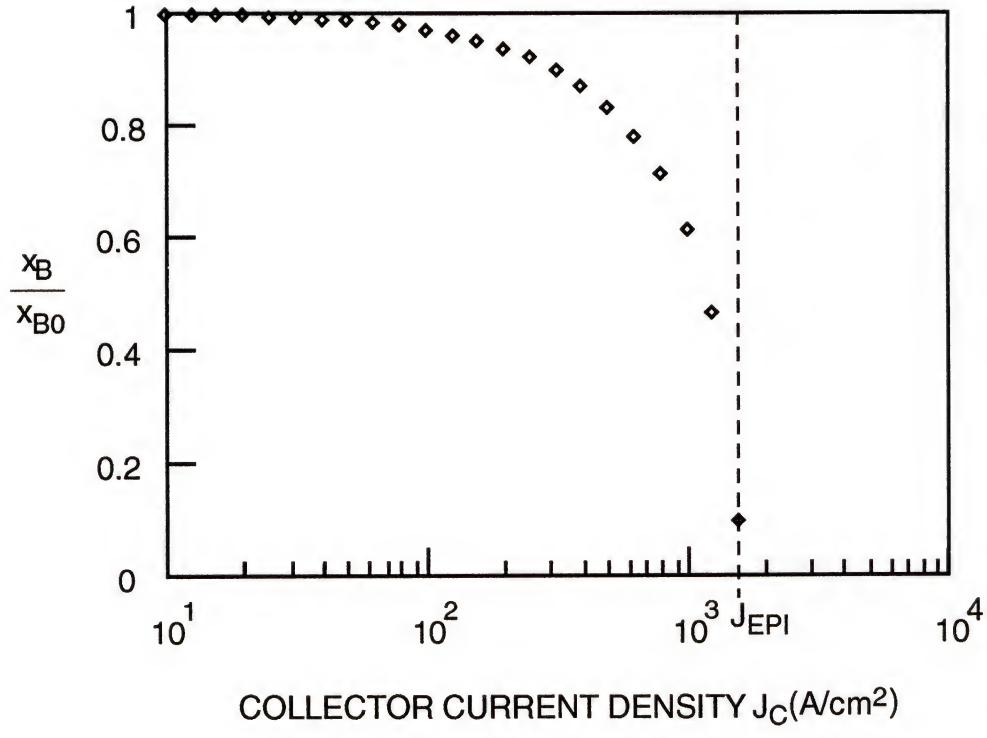


Fig. 2.3. Plot of the ratio x_B/x_{B0} as a function of collector current density J_C for N_A = $1x10^{17}$ cm^{-3} and N_{EPI} = $1x10^{15}$ cm^{-3}.

Similarly, an expression for the Early voltage neglecting velocity saturation can be found from (2.8), (2.9), and (2.10), assuming n_C=0, as

$$V_{A0} \cong \frac{W_{B0}\,(V_{CB}+\phi_C)}{0.5\,x_{B0}} \quad . \tag{2.12}$$

Equation (2.11) is similar in form to (2.1) except for the addition of a third term in the brackets. The last two terms have only a small effect on V_A except for very narrow base-widths. For $J_C > 0.1\ J_{EPI}$, however, (2.11) predicts more significant increases in V_A due to the $(1-J_C/J_{EPI})^{1/2}$ dependence of x_B in the factor $W_B(V_{CB}+\phi_C)/(0.5\,x_B)$ in (2.11).

Equation (2.11) is valid only up to the value of J_C for which the base-collector SCR extends all the way to the buried layer for a given value of V_{CB}, i.e., when $x_C = W_{EPI}$. The upper limit of validity for (2.11) can be found from (2.3) and (2.5) with x_C replaced by W_{EPI}:

$$J_C\,|_{x_C=W_{EPI}} \cong \frac{J_{EPI}}{1+\dfrac{N_A}{N_{EPI}}\left(\dfrac{x_{B0}}{W_{EPI}}\right)^2}\left[1 - \left(\frac{x_{C0}}{W_{EPI}}\right)^2\right] \tag{2.13}$$

where $$x_{C0} = \sqrt{\frac{2\varepsilon_s N_A\,(V_{CB}+\phi_C)}{qN_{EPI}\,(N_{EPI}+N_A)}} \tag{2.14}$$

is the value of x_C obtained from the depletion approximation. For typical BJTs, $(N_A/N_{EPI})(x_{B0}/W_{EPI})^2 << 1$. Thus, when $W_{EPI} >>$

x_{CO}, (2.11) is valid for values of J_C very close to J_{EPI}, so that the effect of carrier-velocity saturation on V_A can be very large.

2.2.2. Case 2

As J_C increases above $J_C|_{x_C = W_{EPI}}$ the collector-base SCR begins to uncover charge in the buried layer. The resulting electron concentration and electric field distribution are shown in Fig. 2.2b. As the current increases above J_{EPI}, the electron concentration in the epi rises above N_{EPI}, leading to field and electron concentration distributions as shown in Fig. 2.2c. Using charge neutrality and solving Poisson's equation over the collector junction SCR, x_B for Case 2 (Figs. 2.2b and 2.2c) can be written as

$$x_B \cong -W_{EPI} + \sqrt{W_{EPI}^2\left[1 + \frac{N_{EPI}}{N_A}\frac{(1 - J_C/J_{EPI})}{(1 + J_C/J_A)}\right] + \frac{2\varepsilon_s(V_{CB} + \phi_C)}{qN_A(1 + J_C/J_A)}} \qquad (2.15)$$

From (2.7), (2.10), and (2.15), the Early voltage V_A can be expressed as

$$V_A \cong \frac{qN_A}{2\varepsilon_s}\left(1 + \frac{J_C}{J_A}\right)\left\{2\left(W_B + \frac{\kappa D_n}{v_s}\right)(W_{EPI} + x_B) + \frac{J_C/J_A}{(1 + J_C/J_A)^2}\left[W_{EPI}^2 + \frac{2\varepsilon_s(V_{CB} + \phi_C)}{qN_A}\right]\right\}. \qquad (2.16)$$

2.2.3. Case 3

Further increments in J_C drive the BJT into non-ohmic quasi-saturation (NOQS). Figure 2.2d shows a typical electron concentration and electric field distribution in NOQS. In NOQS operation, the epi-SCR excess free-electron concentration (n_C - N_{EPI}) is balanced by the heavily doped buried-layer charge. Solution of Poisson's equation leads to

$$x_{EPI} \cong \sqrt{\frac{2\varepsilon_s(V_{CB}+\phi_C)}{qN_{EPI}(J_C/J_{EPI}-1)}} \tag{2.17}$$

where x_{EPI} is the epi-layer SCR width in NOQS operation (Fig. 2.2d). In this case W_B must be redefined as

$$W_B = W_{JC} + W_{EPI} - x_{EPI}. \tag{2.18}$$

Combining this expression with (2.10) and (2.18) yields an Early voltage expression for NOQS mode:

$$V_A = \frac{W_B(V_{CB}+\phi_C)}{0.5\,x_{EPI}}\left[1+\frac{\kappa D_n}{v_s W_B}+\frac{\varepsilon_s v_s (V_{CB}+\phi_C)J_C}{x_{EPI}W_B(J_C-J_{EPI})^2}\right]. \tag{2.19}$$

The upper limit of validity for (2.16) and the lower limit for (2.19) can be found by replacing x_{EPI} with W_{EPI} in (2.17):

$$J_C\big|_{x_{EPI}=W_{EPI}} \cong J_{EPI}\,[1+(x_{C0}/W_{EPI})^2] \tag{2.20}$$

Figure 2.4 shows a plot of the Early voltage using (2.11), (2.16), (2.19), and using the "conventional"

analysis, (2.12), assuming $\kappa = 1$ and $D_n = 17\ cm^2/V{\cdot}s$. In (2.11), (2.16), and (2.19), κ is usually greater than 1, especially for high currents, but this has only a small effect on the resulting Early voltage.

In Fig. 2.4, discrepancies between the conventional and improved models at low currents are as Roulston predicted. The dependence of the SCR boundary position on electron concentration has little effect for low currents. However, for J_C greater than about 0.1 J_{EPI}, this effect becomes significant and V_A increases rapidly as J_C approaches J_{EPI} until the entire epi-layer becomes space-charged (Case 1). As mentioned earlier the potential drop across the QN epi is neglected in (2.11). At low currents this drop is negligible. As the current approaches J_{EPI}, the IR drop becomes more significant, increasing V_A slightly more than predicted by (2.11). However, the resistance of the QN epi itself drops with increasing current due to the reduced width of QN epi. Thus V_A should again be as predicted by (2.11) at the boundary between Case 1 and 2, where the QN epi width reaches zero.

For J_C greater than $J_C|_{x_C = W_{EPI}}$ (Case 2), the Early voltage continues to increase, but at a slower rate than in Case 1. In this case, the heavier buried-layer doping reduces the rate of base push-out with increasing J_C.

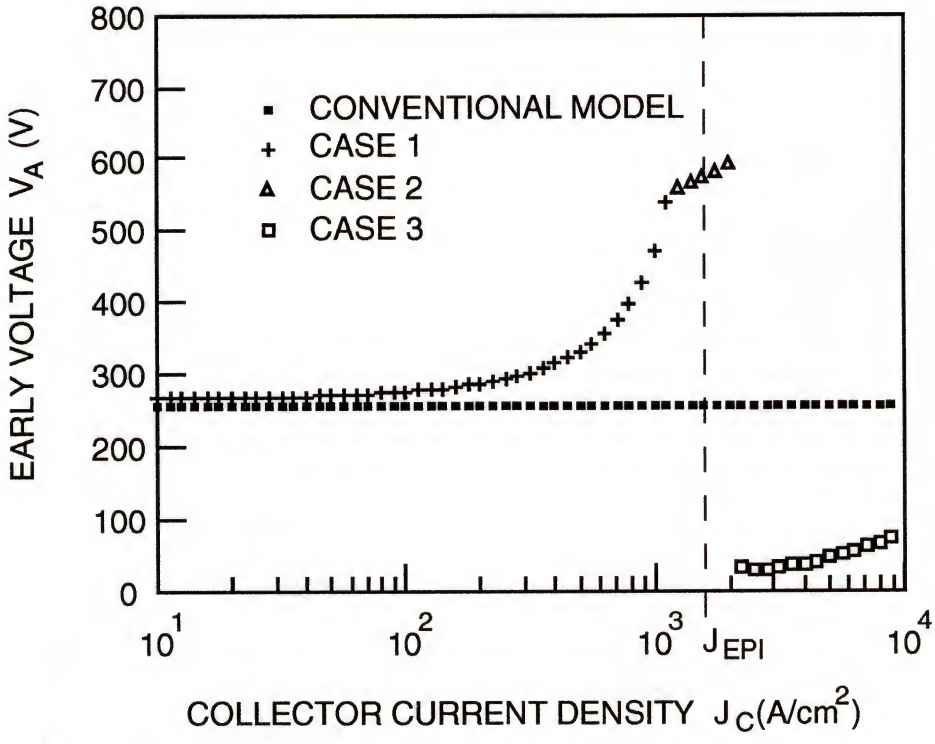


Fig. 2.4. Comparison of conventional and improved V_A models for a transistor with N_A=1x10^{17} cm^{-3}, N_{EPI}=1x10^{15} cm^{-3}, W_{JC}=0.5 μm, W_{EPI}=7 μm, and V_{CB}=10 V as a function of J_C.

As can be seen in the figure, V_A drops at the transition between Case 2 and 3 as the device enters non-ohmic quasi-saturation. As shown, V_A can even drop below the prediction of the conventional model. For $J_C >> J_{EPI}$, V_A becomes proportional to $(J_C)^{1/2}$.

For very high base-collector reverse-biases or very narrow epi thicknesses the epi-layer can become completely depleted even at low currents. For this case, (2.11) is not valid, and (2.16) is valid for any current up to the value of J_C given by (2.20). Equation (2.19) remains valid for NOQS operation. Figure 2.5 shows the Early voltage variation as a function of J_C for $\kappa = 1$, $D_n = 17$ cm^2/V·s, and $V_{CB} = 3$ V for a BJT with a 0.35 μm epi thickness, which is thin enough to allow epi-layer depletion at low currents. As shown in the figure, carrier-velocity saturation can still enhance the output resistance considerably in forward-active mode for $J_C > J_{EPI}$.

2.3. Numerical Simulations

A series of numerical simulations were performed using PISCES-II and MMSPICE to verify that these effects can indeed occur. In PISCES-II the effect of carrier-velocity saturation on Early voltage was studied by performing simulations with and without the field-dependent mobility which models carrier-velocity saturation. Several BJT structures were

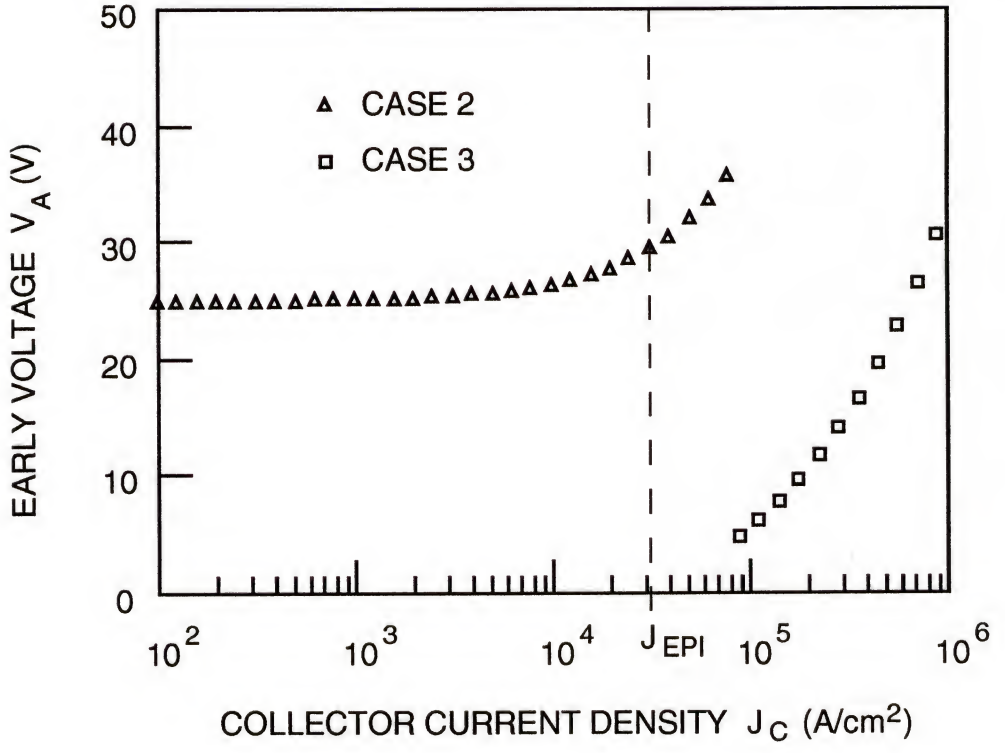


Fig. 2.5. Plot of the improved V_A model for a transistor with $N_A=0.5x10^{18}$ cm^{-3}, $N_{EPI}=1x10^{16}$ cm^{-3}, $W_{JC}=0.1$ μm, $W_{EPI}=0.35$ μm, and $V_{CB}=3$ V. The epi-layer is depleted completely even at low current.

simulated, and the results were consistent with analysis. Figure 2.6 shows a comparison of V_A behavior with and without field-dependent mobility for a transistor with a 1.9 μm thick epi-layer. As can be seen in Fig. 2.6, using the field-dependent mobility model, V_A does increase considerably at high currents, whereas with the field-dependent mobility turned off, V_A increases only slightly (due to emitter-side base-width modulation). In the lower curve (with field-dependent mobility turned off) the Early voltage drops at high currents. This effect can be explained by the onset of ohmic quasi-saturation.

The physics-based BJT model in the circuit simulator MMSPICE [Je89b] includes a charge-based model for advanced high-speed transistors based on one-dimensional analysis of carrier transport. The analysis accounts for high-current effects and impact ionization. The model covers all of the cases considered in the above analyses. Appropriate charge equations are applied to the various regions of the device, allowing for moving regional boundaries. The resulting set of coupled equations require iterative numerical solution. The highly physical basis of this approach allows the software to be used for device simulation as well as for circuit simulation.

We have modified MMSPICE to allow output of internal physical parameters (e.g., x_B, x_C, QN base-width, and QN epi-

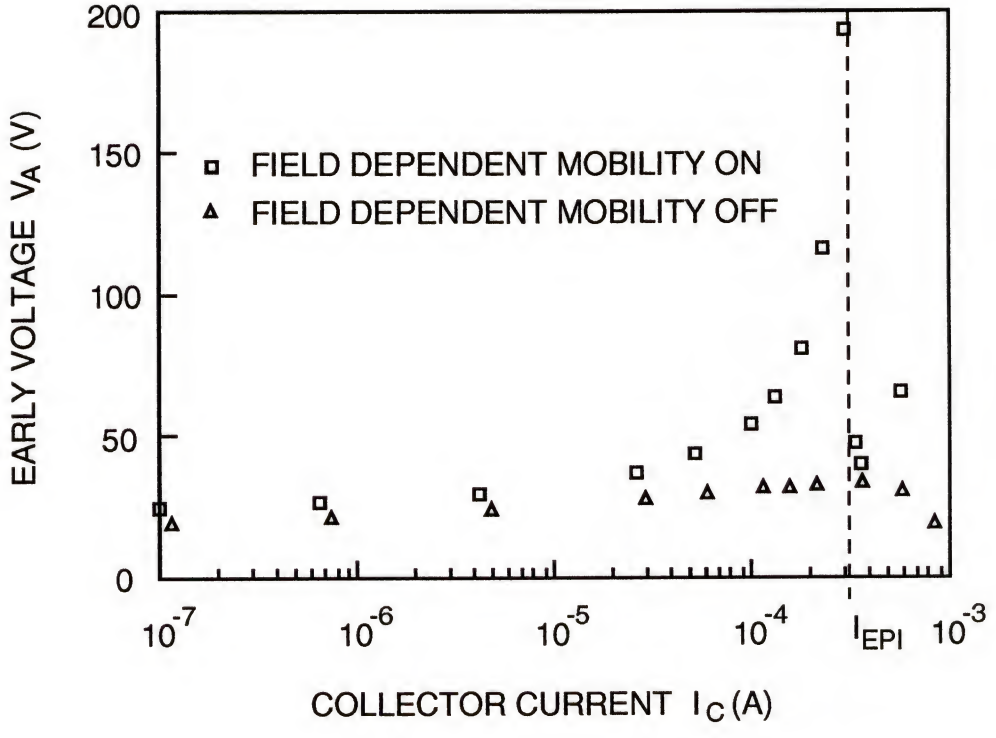


Fig. 2.6. PISCES-II simulation of Early voltage variation as a function of collector current. The device was similar to MOSAIC III BJT [8] except the epi was extended to 1.9 μm. $N_{EPI}=1x10^{16}$ cm^{-3}, $W_{JC}\cong 0.15$ μm, and $V_{CB}=10$ V.

width), in addition to the usual voltage and current outputs. Figure 2.7 shows V_A and x_B variations for a typical analog BJT as a function of I_C for a fixed V_{CE}. As can be seen in the figure, MMSPICE predicts substantial increases in Early voltage due to carrier velocity saturation at high currents. Figure 2.7 also clearly shows that the Early voltage enhancements are associated with variations in x_B.

2.4. Conclusions

The analysis in Section 2.2 demonstrates that BJT output resistance can be raised significantly by carrier-velocity saturation. For typical base-widths, there are usually small increases in output resistance at low currents, but the most important effects are predicted for currents close to the critical value J_{EPI}. These effects are also predicted by PISCES and by the highly physical circuit/device simulator MMSPICE.

However, the predicted peak in V_A cannot be observed in low-frequency measurements. The reason for this inaccuracy is the neglect of the rise in emitter junction temperature caused by an increase in collector junction power dissipation. Thermal feedback tends to severely degrade the output resistance at high currents [Mu64, Me77, Fo91a, Fo91b, Fo91c] which tends to mask the V_A enhancement. With an operating point chosen near J_{EPI} it may be possible to design

high-speed, high-power amplifiers with very high voltage gain if operation is restricted to frequencies high enough to avoid thermal feedback.

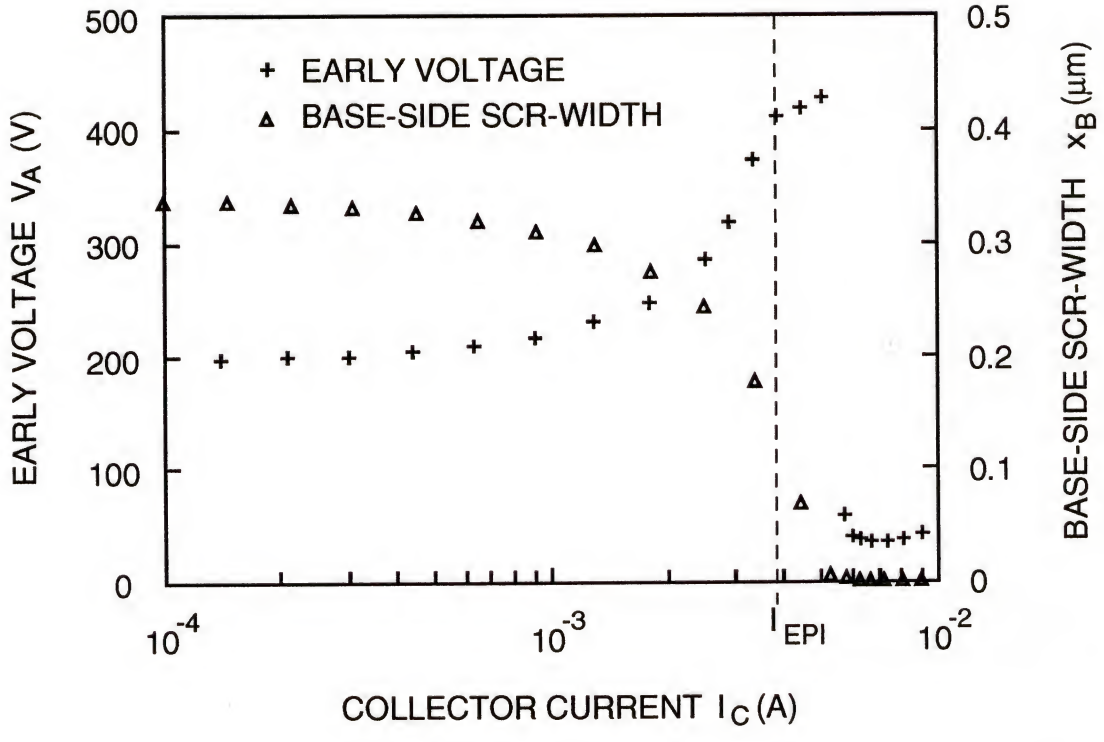


Fig. 2.7. MMSPICE simulation results of Early voltage and SCR penetration into the base as a function of collector current for a typical analog BJT with N_{EPI} = 1x10^{15} cm^{-3}, base-width ≅ 0.8 μm, epi-width ≅ 7 μm, and V_{CB} = 15 V.

CHAPTER 3
SCALABLE SMALL-SIGNAL MODEL FOR BJT SELF-HEATING

3.1. Introduction

It has long been known [Sp58, Me77] that collector-to-emitter thermal feedback can profoundly affect bipolar transistor performance. This effect is becoming more important as current densities and thermal spreading impedances rise with technology scaling. Previously developed thermal impedance models were based in the time domain [Jo70]. It was shown in [De89] that neglecting thermal effects caused errors up to 4% in the predicted delay of a simple BJT logic gate. This chapter will show that local heating causes much larger errors in BJT small-signal parameters. A new model is derived that predicts thermal corrections to BJT small-signal parameters based on a solution of the heat-flow equation in the frequency domain. The model is applicable in principle to any emitter geometry. Polynomial approximations provide simplified expressions for rectangular emitters. These expressions are simple enough for compact modeling in SPICE.

3.2. Model for Thermal Spreading Resistance

Figure 3.1 shows typical measurements of the emitter temperature response to a negative step in collector power. The measurement set-up is shown in Fig. 3.2. In Fig. 3.2, heating current I_H is turned off during brief intervals, during which $V_{BE}(t)$ is measured for small fixed current I_M. Electrical transients dominate the earliest time scale of the measurement, so the temperature is plotted only for $t > 0.4$ μs. This presumably hides a short delay before the emitter temperature begins to respond. The emitter temperature continues to fall for several microseconds. After 100 μs the local-heating transient is complete and the temperature is constant. Then, about 10 ms after the power step, the whole package begins to cool; this transient continues for many minutes. The cooling of the package is controlled by the package thermal impedance, which causes a coupling of all the devices on the chip. Figure 3.3 shows the effect of a negative power step in one transistor's power on the emitter temperature of a nearby transistor. The measurement setup is shown in Fig. 3.4. Only the longer time-scale transient, controlled by the package thermal impedance, is visible. Consistent with the model of [Jo70], thermal spreading impedance falls off strongly with distance, so that for modest powers, even adjacent devices are thermally decoupled

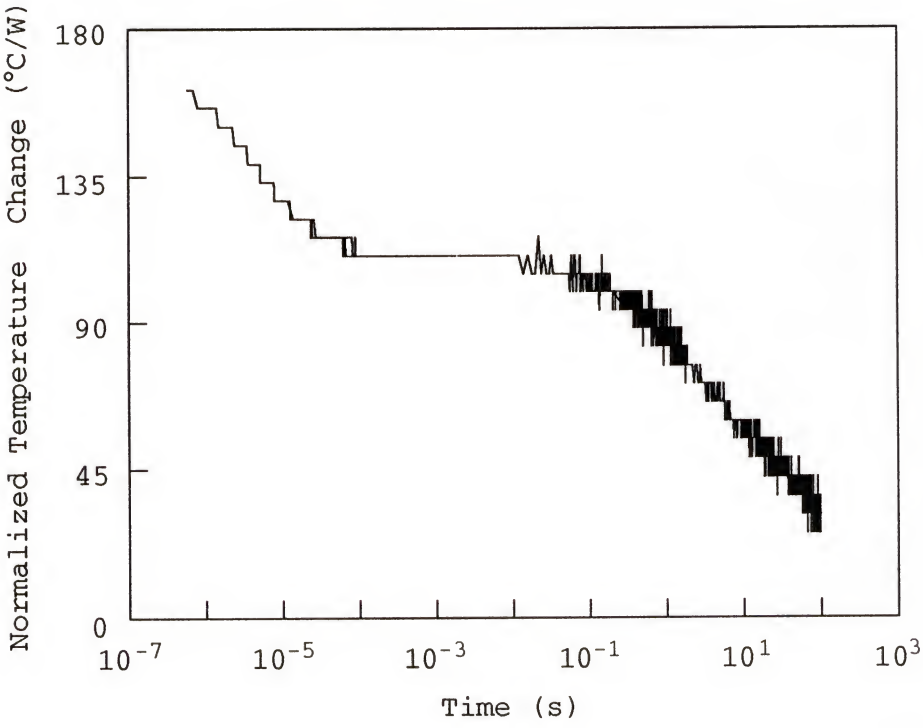


Fig. 3.1. Measured response to negative collector-power step. Results of several measurements are superimposed.

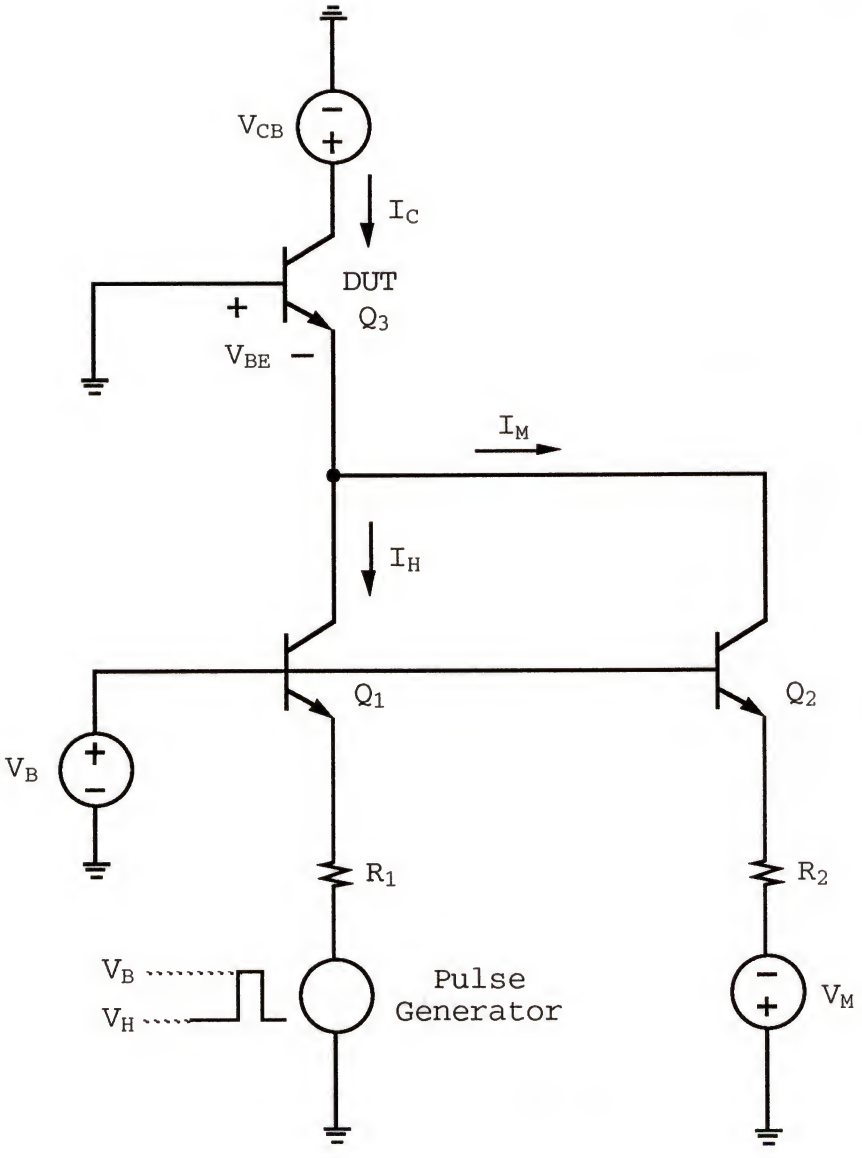


Fig. 3.2. Experimental setup for self-heating. Heating current I_H is turned off during brief intervals, during which $V_{BE}(t)$ is measured for small fixed current I_M.

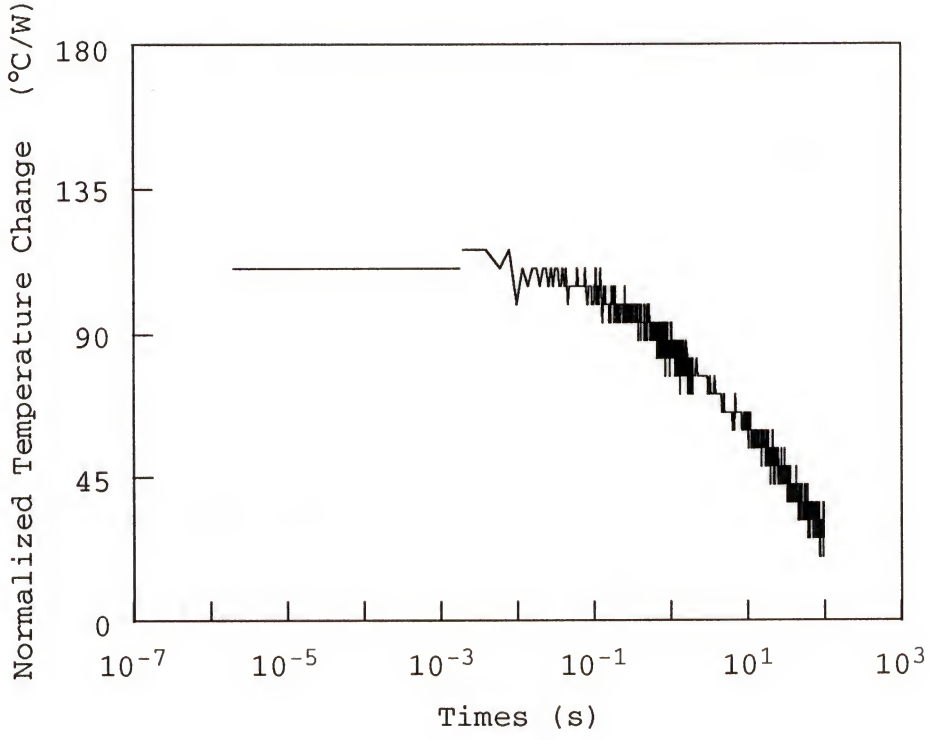


Fig. 3.3. Measured response of transistor heated by nearby transistor.

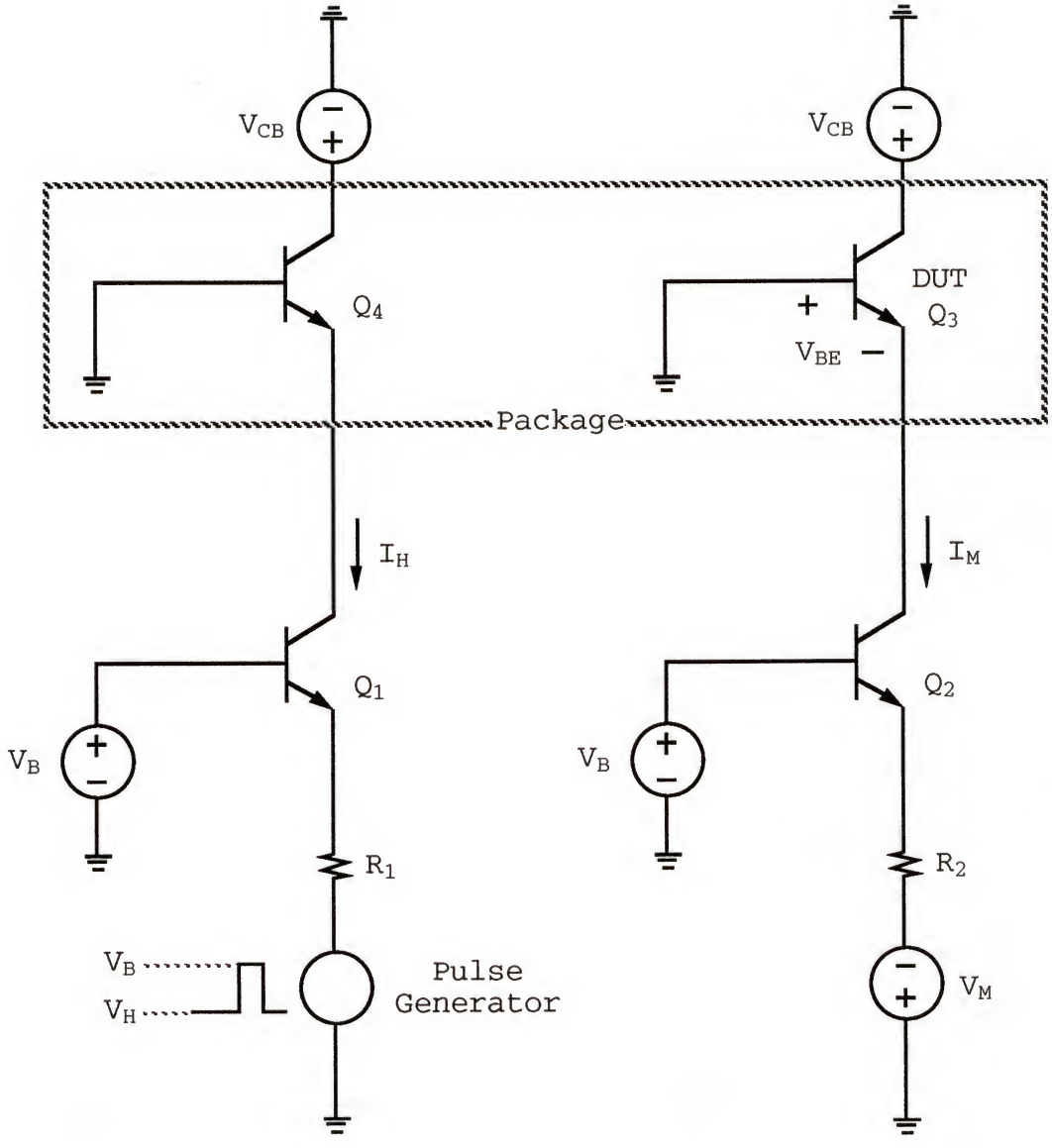


Fig. 3.4. Setup for measuring thermal response of transistor heated by nearby transistor.

over the shorter time scale. This work is concerned primarily with local heating effects.

If a unit impulse of heat occurs instantaneously at some point in a uniform medium, the temperature rise at a point a distance r away is given by [Ca62]

$$T(r,t) = \frac{1}{8\rho c(\pi\kappa t)^{3/2}} \exp\left(\frac{-r^2}{4\kappa t}\right) \qquad (3.1)$$

where ρ is the density, c is the specific heat, and κ is thermal diffusivity. An identical image source can be added to create an adiabatic surface at the plane equidistant between the sources. In their time-domain derivation, Joy and Schlig [Jo70] integrated (3.1) over the volume of the collector space-charge region (SCR) and its image. The result is a closed-form expression for the impulse response at any point in the semiconductor. The time integral of this expression gives the response to a unit step in the collector power. The integral is complicated and does not have a closed-form solution. The step response at $t = \infty$ gives the thermal resistance R_{th}, defined as the total temperature rise caused by a unit step in the collector power.

Note that (3.1) is equivalent to thermal resistance which is defined as the increments in temperature for a given incremental unit power. Therefore, a frequency-domain model

can be derived by first taking the Laplace transform of (3.1), yielding

$$Z_{th}(r,s) = \frac{1}{4\pi Kr} \exp(-r\sqrt{s/\kappa}) \tag{3.2}$$

where $K = \rho\kappa c$ is thermal conductivity. For $s = 0$, this equation gives R_{th} for the point source. R_{th} can be integrated over the collector SCR and its image to yield the thermal resistance at any point $r' = (x', y', z')$ for arbitrary geometry. In rectangular coordinates

$$R_{th}(W,L,H,D,r') = \frac{1}{4\pi K} \int_0^W \int_0^L \left[\int_D^{D+H} \frac{dxdydz}{r} + \int_D^{-(D+H)} \frac{dxdydz}{r} \right] \tag{3.3}$$

where $r = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$, W and L are the width and length of the emitter, D is the depth of the base-collector junction, and H is the SCR width. This integral is not solvable in closed form, so least-squares fitting was used to generate the following approximation to (3.3), in which r' is taken as the point at the surface above a corner of the emitter:

$$R_{th} = 1/2\pi K r_{eff} \tag{3.4a}$$

$$r_{eff} = 2\sqrt{WL}\, f_1 f_2 \tag{3.4b}$$

$$f_1(d,h) = (0.058d+0.14)h+0.34d+0.28 \tag{3.4c}$$

$$f_2(a) = 0.98+0.043a-(6.9\cdot10^{-4})a^2+(3.9\cdot10^{-6})a^3 \qquad (3.4d)$$

where $d = D/\sqrt{WL}$, $h = H/\sqrt{WL}$, and $a = W/L$. In (3.4a), the 2 instead of 4 in the numetrator is given by the fact that both collector-junction SCR and image are considered as a point-source of its own. The approximation agrees with numerical solution of (3.3) to within 5% for all reasonable geometries. Figures 3.5 and 3.6 show how f_1 and f_2 vary with device geometry. Figure 3.7 is a contour plot of predicted R_{th} variation with position within the emitter for a transistor with W = 64 μm, L = 48 μm, H = 4 μm, and D = 3 μm. Only one quadrant is shown and the center of the emitter is at the upper left. To find R_{th} for any other point r' at the surface above the emitter, the transistor can be broken into four smaller rectangular transistors, meeting at r'. Equation (3.4) can be used with modified values for W and L to compute R_{th} for each quadrant. The overall value for R_{th} can be found using

$$R_{th} = \frac{1}{WL}\sum_{i=1}^{4} R_{thi}W_iL_i \qquad (3.5)$$

Table 1 gives calculated R_{th} values for r' located in the center of the emitter and at the corner for several transistor geometries. Values for H were computed using the step-junction approximation. We believe this is the first

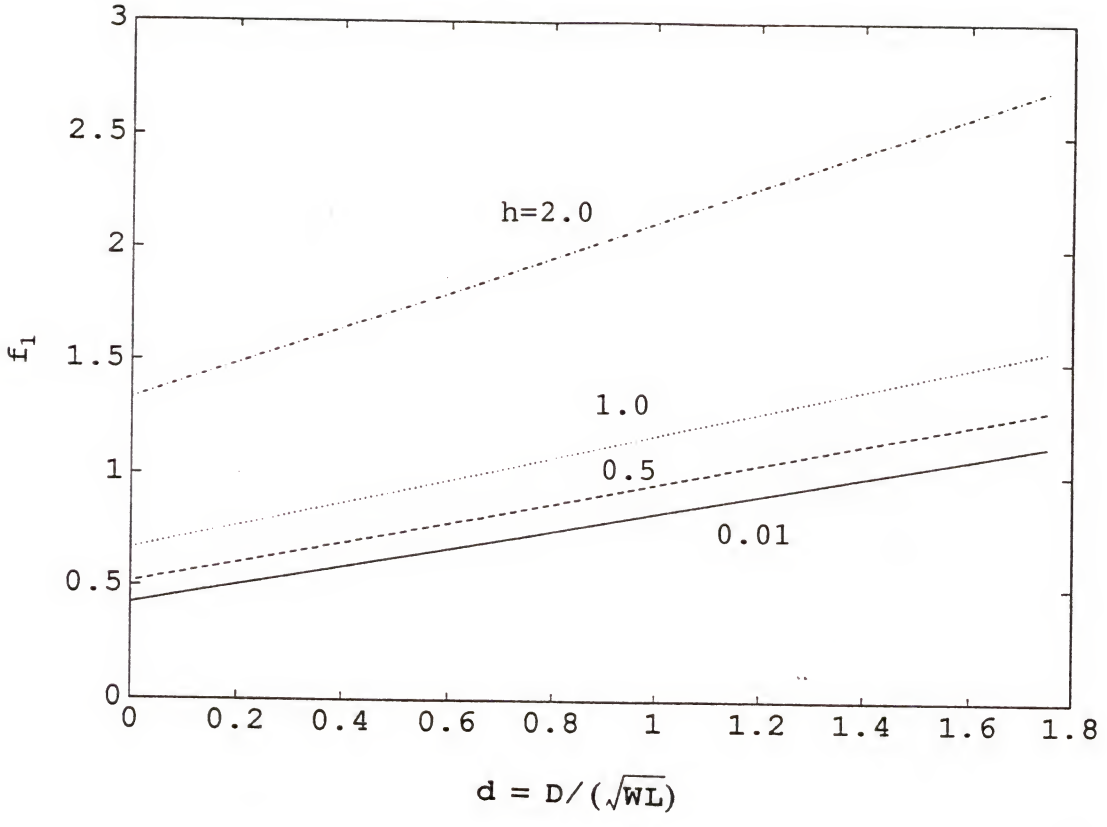


Fig. 3.5. Variation of function $f_1(d,h)$ with $d = D/\sqrt{WL}$, with $h = H/\sqrt{WL}$ as a parameter.

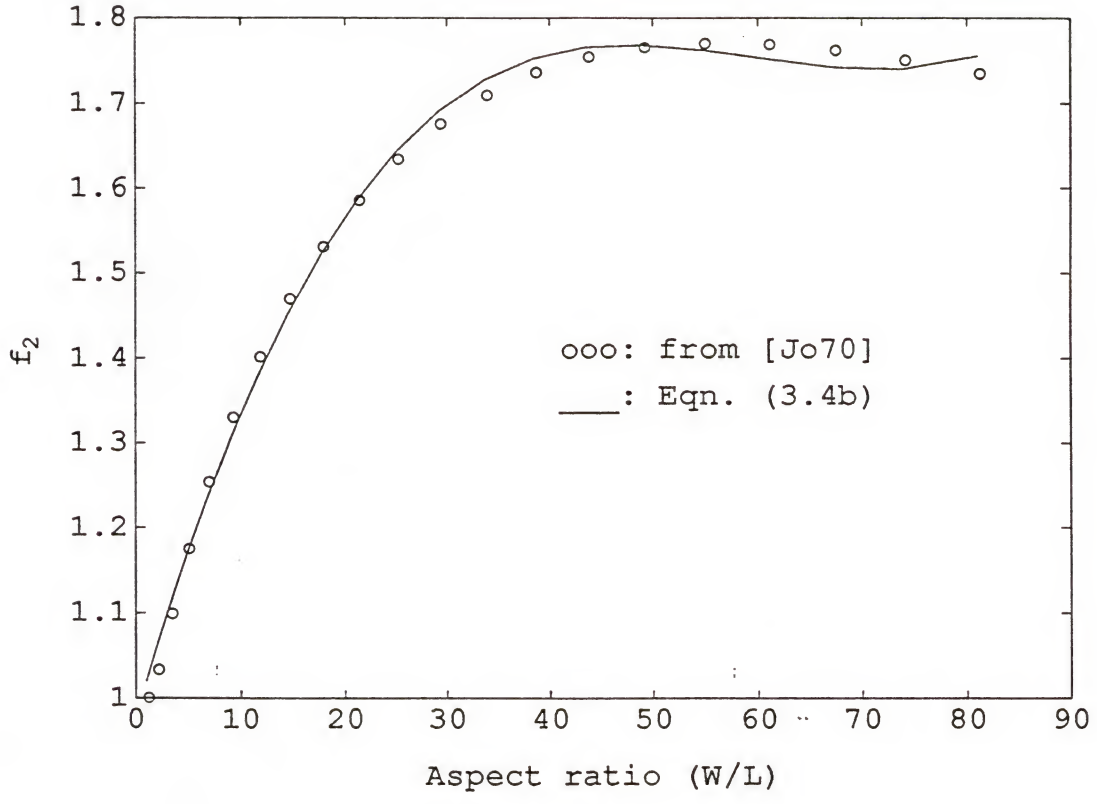


Fig. 3.6. Variation of function $f_2(a)$ with aspect ratio a = W/L.

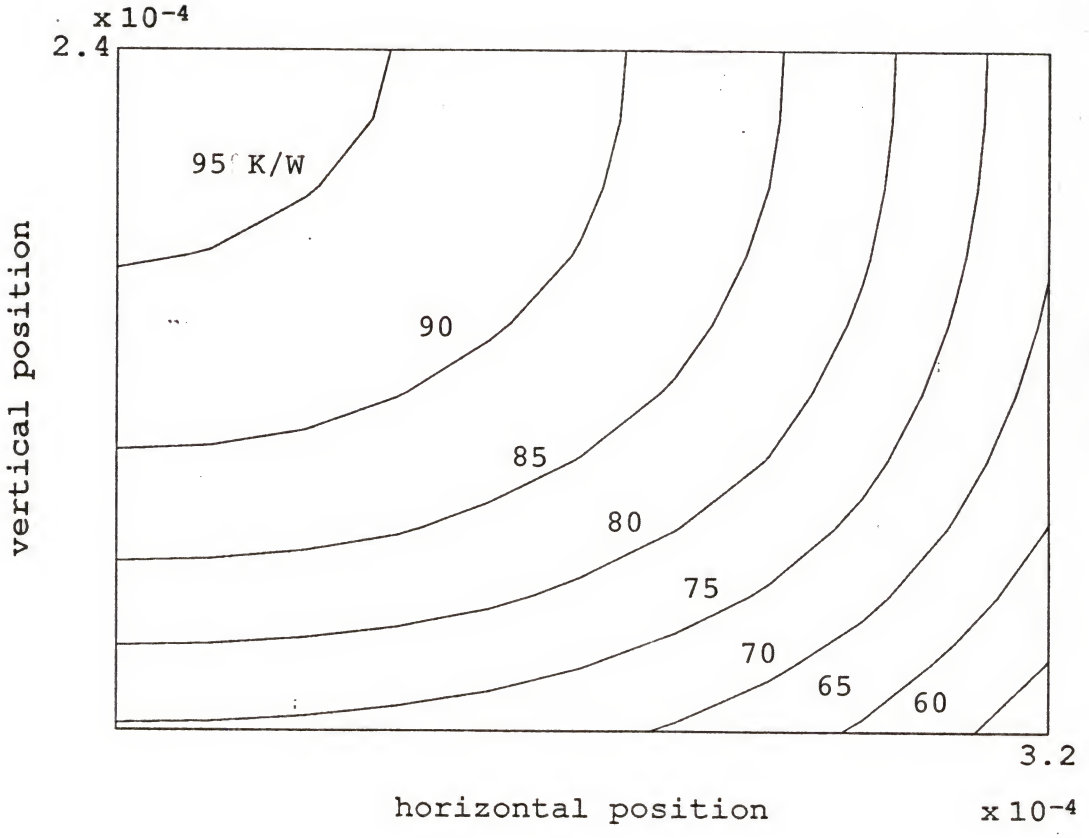


Fig. 3.7. Contour plot of thermal resistance at the surface versus position within emitter for a transistor with W = 32 μm, L = 24 μm, H = 4 μm, D = 3 μm. Only on quadrant is shown and the center of the emitter is at the upper left.

compact model which can predict the variation of R_{th} across the emitter and the variation with D and H.

3.3. Small-Signal Model

Mueller [Mu64] showed that BJT common-emitter y-parameters can be corrected for thermal feedback using

$$Y_{mn} = \frac{Y_{mnE} + D_m Z_{th} I_m I_n}{1 - D_m Z_{th} P} \tag{3.6}$$

where m or n = 1 for the base or 2 for the collector, Y_{mnE} is the uncorrected electrical Y parameter, Z_{th} is the thermal impedance, and P is the dissipated power. D_2 is the fractional temperature coefficient of the collector current, (typically about 7 %/K) and D_1 is that of the base current, equal to $D_2 - D\beta$, where $D\beta$ is the fractional temperature coefficient of β, (typically around 0.7 %/K). Equation (3.6) can be used with $Z_{th} = R_{th}$ to thermally correct the DC small-signal parameter g_{mn}.

Figure 3.8 compares measured and simulated values for the normalized output resistance parameter $V_A = I_C/g_{22} - V_{CE}$ for a silicon BJT with 23 x 23 $(\mu m)^2$ emitter, collector junction depth of 3.0 μm, and epi-doping of 10^{15} /cm^3. The value for H in (3.4) was 2.8 μm, calculated assuming a step junction with V_{CE} = 6 V, which led to R_{th} = 70 °C/W. The simulation used $g_{22E} = I_C/V_{AE}$, where V_{AE} = 350 V is the

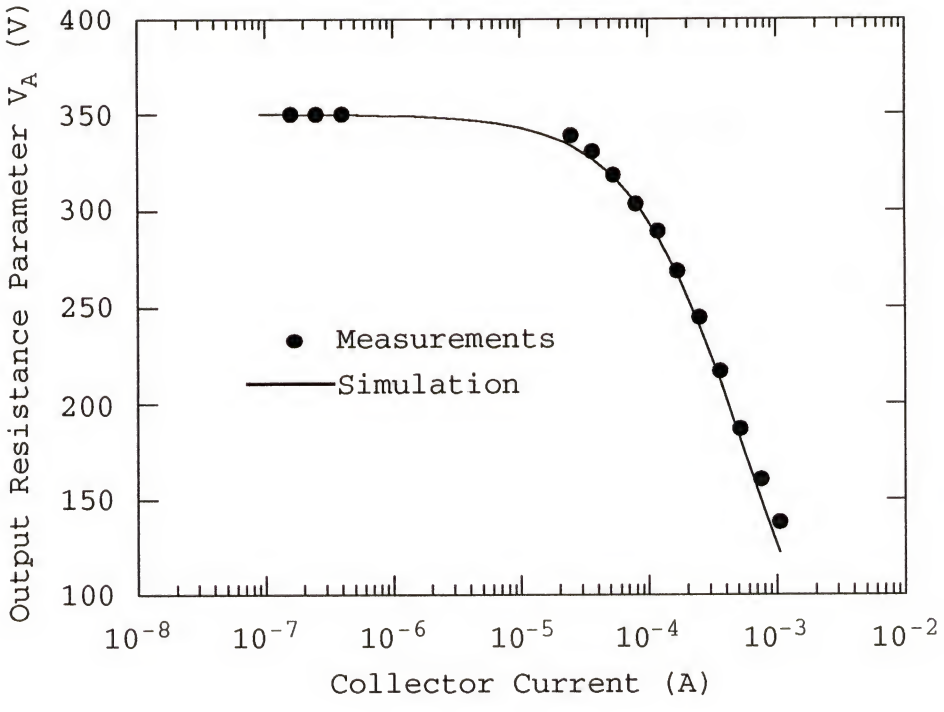


Fig. 3.8. Measured and simulated normalized output resistance $V_A = I_C/g_{22} - V_{CE}$. Simulations are based on (3.4) and (3.6).

Table 1: Calculated and measured thermal resistance for several transistor geometries. Calculations were based on (3.4) and (3.5). Measured values give best fits to output resistance and reverse transconductance using (3.6) and include an estimate of experimental error.

Type	W (μm)	L (μm)	R_{th} at corner	R_{th} at center	Measured R_{th}
a	23	23	69	115	74±5
	83	23	36	65	41±4
	12	12	112	170	118±10
b	4	2	492	778	480+20
	8	2	348	582	390±5
	10	2	306	520	350±20
	10	4	251	443	290±10
c	7.2	1.2	467	814	435±20

a: D = 3 μm, N_{epi} = $1x10^{15}$ cm^{-3}, V_{CE} = 6 V, H = 2.83 μm.

b: D = 0.4 μm, N_{epi} = $7x10^{15}$ cm^{-3}, V_{CE} = 3 V, H = 0.85 μm.

c: D = 0.25 μm, N_{epi} = $2x10^{16}$ cm^{-3}, V_{CE} = 2 V, H = 0.35 μm.

electrical Early voltage, measured at low currents. The data were corrected to account for package thermal resistance of 110 °C/W, measured using the plateau near 1 ms shown in Fig. 3.1. Note that most circuit models use a constant value (the Early voltage) for this parameter. Measured g_{22} or g_{12} variations can be used in (3.6) to extract a best-fit value for R_{th}. The last column in Table 1 compares R_{th} values found by these methods to calculated values. This method will be discussed in Chapter 4. Note that the measured values match well with the calculated R_{th} for the corner of the emitter. This result appears to be largely coincidental, but it is fortunate, as it means that a single effective temperature can be assigned to the emitter using (3.4) with good accuracy.

The other DC small-signal parameter which is strongly affected by thermal feedback is g_{12}. Electrical models predict only a small negative value for this parameter, caused by Early-effect modulation of recombination in the quasi-neutral base. This effect is normally neglected. Equation (3.6) predicts a substantial positive value for g_{12} approaching g_{22}/β with increasing I_C. This can have a substantial effect on circuit performance, as the input and output impedances become strongly dependent on the loading at the opposite port. Note that (3.6) shows that errors in small-signal

parameters due to neglecting thermal feedback can occur without large DC power dissipation.

3.4. AC Modeling

Equation (3.2) can be used to model thermal spreading impedance above DC if the r_{eff} from (3.4) is substituted for r in (3.2). The physical significance of r_{eff} is an effective distance between the collector-junction SCR and the point at the surface above a corner of the emitter. The results are compared to the single-pole model [De89, Vo89, Ye91] and the inverse Fourier transform of Joy & Schlig's predicted impulse response in Fig. 3.9. For this case (the same device as in Fig. 3.1) the magnitude of the point-source model differs by less than 6% compared to the Fourier transform of the full time-domain model for frequencies up to f_{pk} = 350 kHz, defined as $1/(2\pi t_{pk})$, where $t_{pk} = r_{eff}^2/(6K)$ = 0.45 μs is the time at which the impulse response in (3.1) reaches its maximum. At this frequency $|Z_{th}|$ has dropped to 18% of R_{th}. Note that the commonly used one-pole approximation predicts most of the drop in $|Z_{th}|$ to occur over one decade of frequency, whereas the $|Z_{th}|$ in Fig. 3.9 varies over three decades of frequency.

The phase response is less accurate, differing by more than 30 degrees for frequencies greater than f_{pk}. These phase errors are easily understood: the model represents the transistor as a point-source collector and a point emitter,

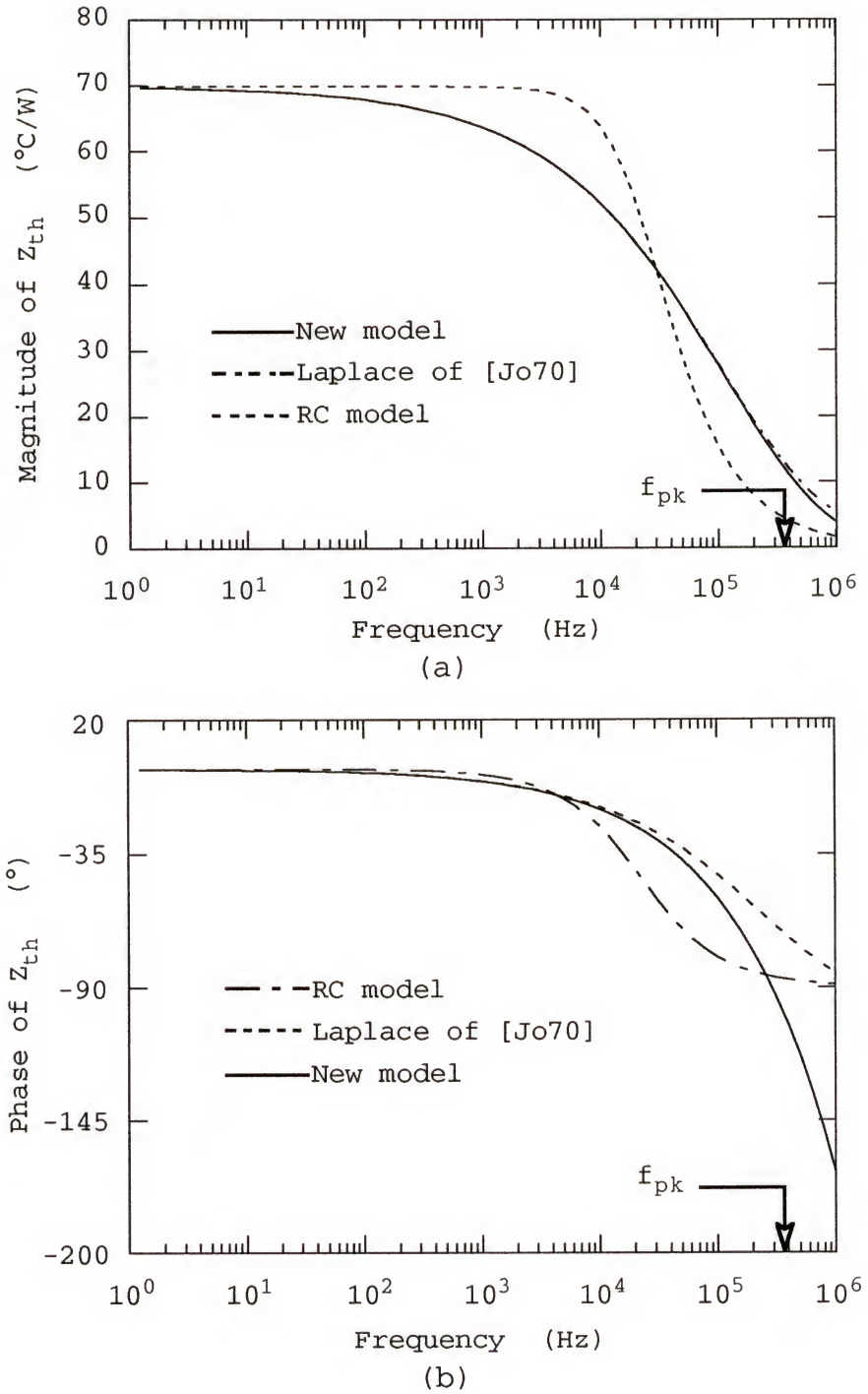


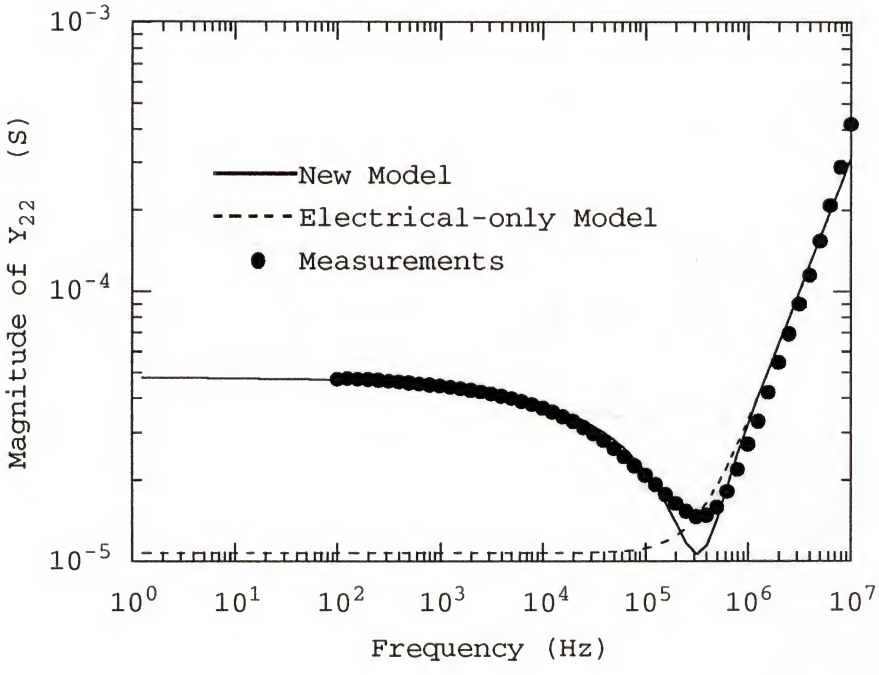
Fig. 3.9. Magnitude and phase of thermal spreading impedance Z_{th}. Solid lines were computed using (3.2) and (3.4).

separated by an r_{eff} which is inevitably greater than the base width. Thus, the model overestimates the initial delay. However, since they coincide with decreasing $|Z_{th}|$, these phase errors cause relatively little error in the y-parameters. Figure 3.10 shows typical measured and simulated values for Y_{22}, both magnitude and phase for V_{BE} = 0.8 V, V_{CE} = 6 V, and I_C = 3.74 mA. As can be seen in the figure, simulations show good agreement with measurements despite the inaccuracy in the phase of the model. Figures 3.9 and 3.10 prove the importance of correctly doing thermal modeling for BJT output conductance.

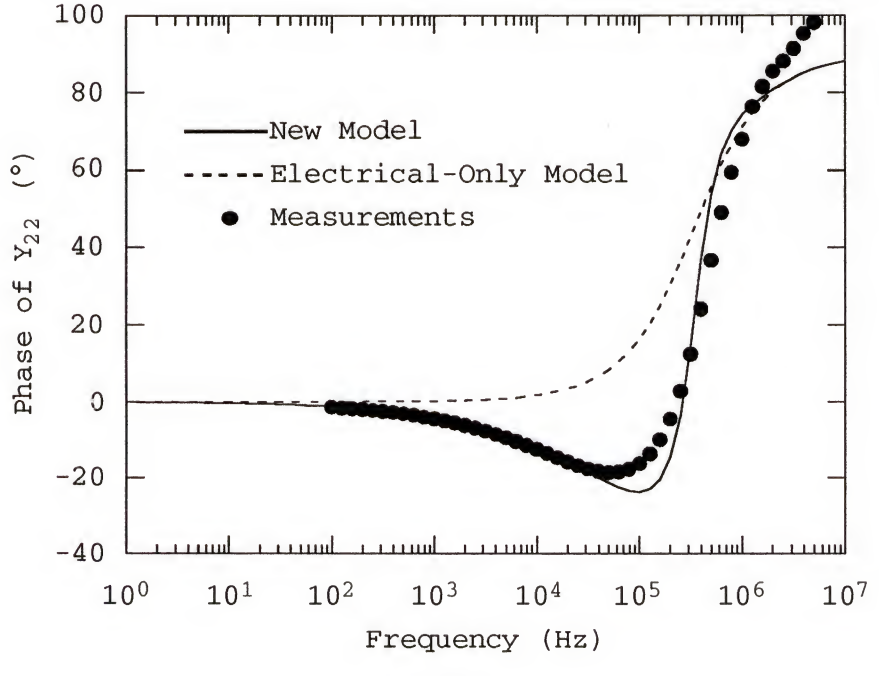
3.5. Conclusion

The model described above provides a simple and practical way to improve the accuracy of BJT circuit simulation. Polynomial approximations are used to express the variation of R_{th} across the emitter; the results are equivalent to the infinite integral in [Jo70]. Measurements show that a single value of R_{th} can represent typical BJT thermal behavior at DC. A frequency-domain solution for Z_{th} is presented which is more accurate than the commonly used single-pole thermal-impedance model.

Considering the significant effects which thermal feedback has on BJT behavior, it is interesting to speculate why thermal effects are not yet routinely included in circuit simulators, in spite of prior work on including such effects,



(a)



(b)

Fig. 3.10. Measured and simulated Y_{22} for V_{BE} = 0.8 V, V_{CE} = 6 V, and I_C = 3.74 mA. (a) Magnitude. (b) Phase.

[Fu76, Vo89]. One contributing factor may be the willingness of many circuit designers to accept empirical models -- model parameters can be adjusted to account for many thermal effects. Figure 3.11 demonstrates that even this empirical approach can lead to significant errors. Extractions of SPICE parameter VAF were done under several different conditions. Differences greater than 30% were found depending on HP4145 integration time, and whether or not a fan was blowing on the package to change its effective thermal impedance. This should give pause to the analog IC designer -- such thermal effects could cause the gain of a two-stage op amp to vary more than 2 to 1 depending on package type and mounting conditions.

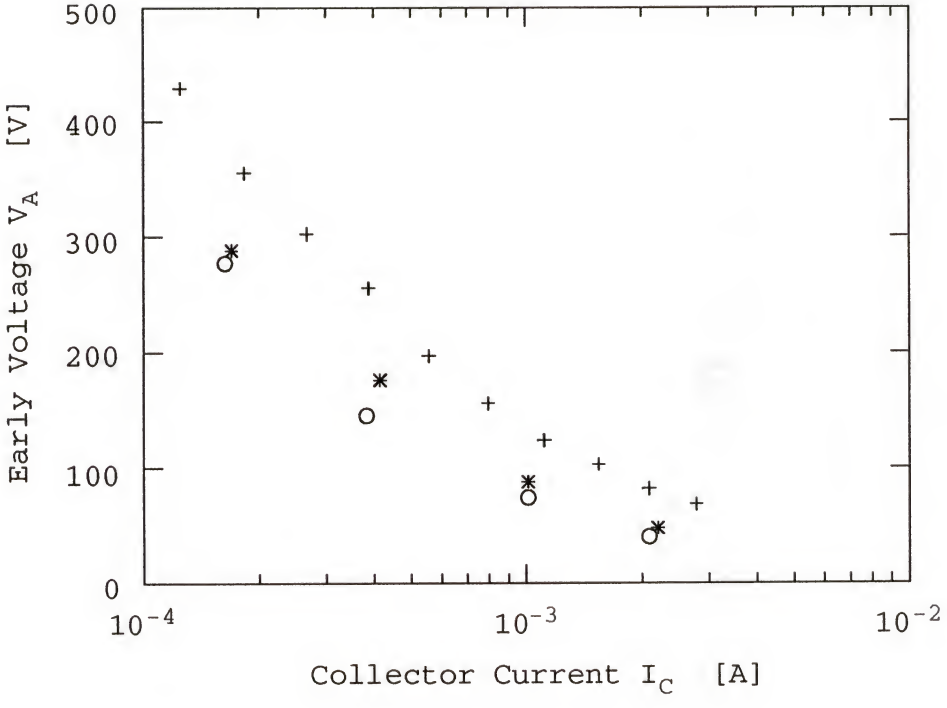


Fig. 3.11. V_{AF} values optimized for various currents, extracted under different thermal conditions: + using an HP-4145 at minimum delay between points; *: static measurement with small fan blowing on the 14-pin DIP; o: static measurement without fan.

CHAPTER 4
PARAMETER EXTRACTION FOR BIPOLAR CIRCUIT MODELING

4.1 Introduction

Thermal feedback from collector power to emitter temperature has been shown to significantly affect BJT performance [Mu64, Jo70, Me77, De89]. Recently several workers [Vo89, Ye91] have demonstrated ways to model thermal feedback in circuit simulation assuming thermal model parameters are available. This chapter introduces simple ways to extract thermal parameters needed for such simulations.

BJT thermal behavior is dominated by two components, characterized by the thermal spreading impedance and the chip/package thermal impedance. It is important to extract these components separately, as the package thermal impedance causes a coupling between transistors on the same chip, and thermal spreading impedance usually does not. Standard methods for extracting thermal impedance are adapted for characterizing power transistors [Oe76]. They are poorly suited for measuring thermal spreading impedance, which becomes increasingly important with decreasing transistor size.

4.2. Thermal Impedance Extraction Methods

The standard electrical method for thermal impedance measurement is the "emitter-only switching" method. In this technique, a large heating current is first applied to establish thermal equilibrium. This current is then turned off and the temperature is inferred from the resulting transient response of V_{BE}, measured at a small fixed emitter current. The initial response is controlled by both electrical and thermal spreading-impedance effects. This makes it difficult to determine the original temperature, and prevents accurate extraction of the thermal spreading resistance R_{th}. The method is still useful, however, for finding package thermal impedance. For small transistors, the transient due to thermal spreading impedance is typically completed long before the thermal transient of the package begins. The package's transient typically begins a few milliseconds after the power step and lasts for up to several minutes [Fo91a]. The package thermal resistance can be found by dividing the total change in temperature after 1 ms by the size of the power step.

4.2.1. Using output resistance as a temperature sensitive parameter

The following technique can be used to extract thermal spreading resistance. It is easy to show [Mu64] that for small changes in the base and collector voltages, the change in I_C can be written as

$$\Delta I_C \cong g_{21E}\Delta V_{BE} + g_{22E}\Delta V_{CE} + I_C D_C R_{th}(I_C\Delta V_{CE} + V_{CE}\Delta I_C) \quad (4.1)$$

where g_{12E} and g_{22E} are common-emitter small-signal two-port parameters neglecting thermal effects, and D_C is the fractional temperature coefficient of collector current ($D_C = I_C^{-1}\ \partial I_C/\partial T$). With ΔV_{BE} set to zero, (4.1) can be solved for $D_C R_{th}$, which gives

$$D_C R_{th} \cong \frac{\Delta I_C/I_C - \Delta V_{CE}/(V_A + V_{CE})}{I_C\Delta V_{CE} + V_{CE}\Delta I_C} \quad (4.2)$$

where g_{22E} is expressed as $I_C/(V_A + V_{CE})$, where V_A is the Early voltage.

The extraction procedure is as follows: V_{CE} is set to some value in the forward-active region. For a series of fixed values of V_{BE}, V_{CE} is varied slightly and the resulting ΔI_C is measured. Since thermal effects are negligible for low currents, a constant value for V_A can be extracted at low V_{BE} values using

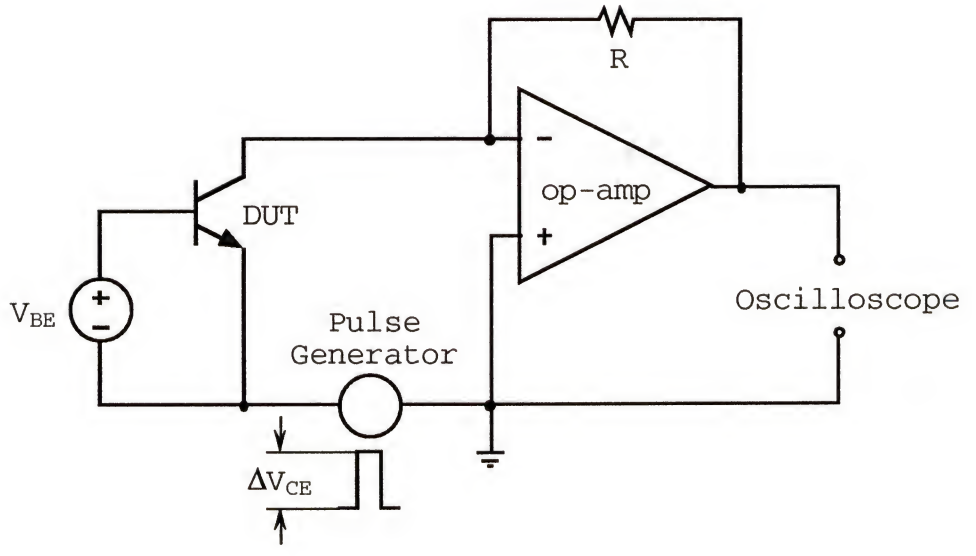
$$V_A \cong (\Delta V_{CE}/\Delta I_C)\ I_C - V_{CE} \quad (4.3)$$

At higher currents, thermal effects become significant, and (4.2) can be used to extract the product $D_C R_{th}$. For moderate currents, D_C can be approximated as [Me77]

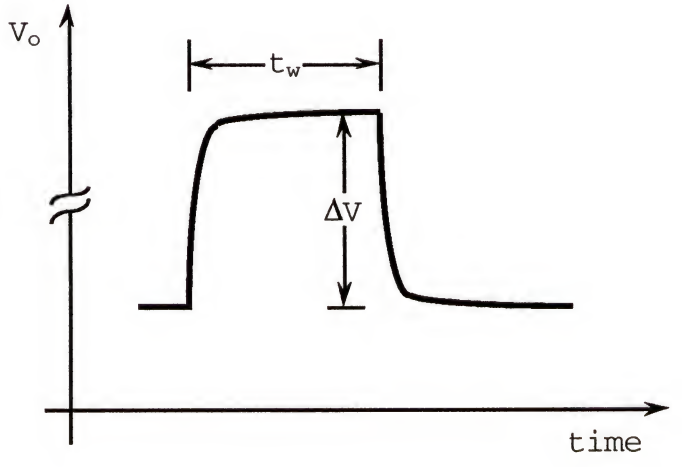
$$D_C = \frac{1}{T} \left[\frac{V_{GO} - V_{BE}}{V_T} + \eta \right), \qquad (4.4)$$

where T is absolute temperature, V_T is the thermal voltage, V_{GO} for silicon is 1.206 V, and $\eta \cong 2$ (related to the doping level). Equation (4.4) becomes inaccurate at very high currents, so it is important not to use such high currents in the extraction.

It is important to carefully consider the timing of these measurements to avoid confusing thermal spreading and package thermal impedances. An HP-4145 Semiconductor Parameter Analyzer set to its minimum integration time requires at least 20 ms to measure a pair of (V_{CE}, I_C) data points, during which time substantial change in package temperature can occur. Longer delays between data points lead to larger errors. (For this reason faster equipment is needed to extract thermal spreading resistance alone.) A pulse generator, a current-to-voltage converter and an oscilloscope are generally adequate. Figure 4.1 show the measurement set-up using such equipment. In the figure, the pulse generator varies the collector current of the device for a given V_{BE}. Variations in collector current are amplified by the op-amp



(a)



(b)

Fig. 4.1 Pulse measurement. (a) Experimental setup to measure R_{th} using a pulse generator, a current-to-voltage converter, and an oscilloscope. (b) A typical shape of measured output.

and measured with an oscilloscope. Variation in collector current is given by $\Delta I_C = \Delta V/R$, where ΔV is the peak-to-peak voltage of op-amp output voltage and R is the op amp's feedback resistor. In Fig. 4.1, the pulse width t_w has to be short enough to avoid any error caused by package thermal impedance.

Another method which avoids the effect of package thermal impedance involves measuring the output admittance Y_{22} using an impedance analyzer. As previously noted, for typical small transistors there is a period after a power step during which the thermal step response is constant for some time. This implies that there is also a range of frequencies over which the thermal frequency response is constant. In the middle of this frequency range Y_{22} is also constant for a given current. Its real part can be denoted g_{22}. At low currents, thermal effects are negligible and $g_{22} = g_{22E} = I_C/(V_A + V_{CE})$, which can be used to find V_A. With increasing current, g_{22} rises above $I_C/(V_A + V_{CE})$, and R_{th} can be extracted using

$$R_{th} = \frac{1}{D_C} \frac{g_{22}/I_C - 1/(V_A + V_{CE})}{I_C - g_{22}V_{CE}} \qquad (4.5)$$

This method generally gives results consistent with those found with the previous method. Figure 4.2 shows a typical

plot of R_{th} extracted at various collector currents based on the above methods.

4.2.2. Using reverse transconductance as a temperature sensitive parameter

In the characterization method described above, the Early voltage, V_A has to be supplied to extract R_{th} from a measured g_{22}. For typical analog devices, the effect of emitter-side base-width modulation ("late effect") is usually negligible so that the Early voltage stays roughly constant with variation in base-emitter voltage for a given collector-base voltage. In this case, the Early voltage measured at low currents can be used, with little error, to extract R_{th} from the g_{22} measured at other currents. However, for typical advanced BJTs, late effects are far more significant because of the narrower base width, so that the extracted Early voltage varies considerably as a function of base-emitter voltage. In this case, using the Early voltage measured at low current can leads to significant error in extracted R_{th}. Extracting Early voltage at high currents is not easy because of the thermal effects. It may still be possible to extract R_{th} using the previous method, but can be cumbersome, and not

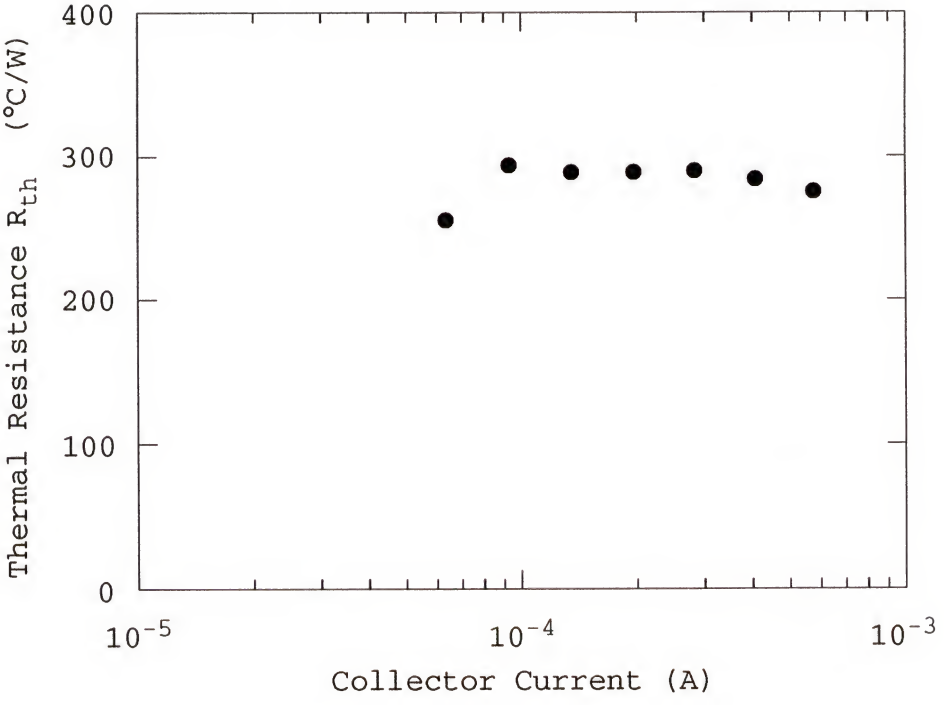


Fig. 4.2 A typical plot of R_{th} extracted at various I_C.

very precise. A better method for this case uses the reverse transconductance g_{12} as the temperature-sensitive parameter. Recently, Reisch [Re92] also have independently suggested this method that will be described in the following.

An equation comparable to (4.1) gives the change in base current as

$$\Delta I_B \cong g_{11E}\Delta V_{BE} + g_{12E}\Delta V_{CE} + I_B D_B R_{th}(I_C\Delta V_{CE} + V_{CE}\Delta I_C) \quad (4.6)$$

where D_B is the fractional temperature coefficient of base-current I_B. For the typical advanced devices, g_{12E} is negligible. With ΔV_{BE} set to 0, (4.6) can be solved for $D_B R_{th}$, which gives

$$D_B R_{th} \cong \frac{\Delta I_B}{I_B(I_C\Delta V_{CE} + V_{CE}\Delta I_C)} \quad (4.7)$$

With similar precautions, the equipment described in the previous extraction method can be applied for this method. For moderate currents, assuming base currents are dominated by back injection into the emitter [Je91b], D_B can be approximated as [Sl76, Sz81]

$$D_B \cong \frac{1}{T}\left[\frac{V_{GO} - V_{BE} - \Delta V_G}{V_T} + \eta\right] \quad (4.8)$$

where ΔV_G is the bandgap narrowing which is a function of the impurity doping concentration in the emitter. The accuracy of the extracted R_{th} depends on the accuracy of ΔV_G and η.

However, the errors are not significant as these parameters are typically small compared to the other terms. Equation (4.8) becomes inaccurate at very high currents due to the base and emitter parasitic resistances.

Note that output resistance should always be extracted with the base voltage fixed, not the base current. Since g_{12E} is extremely small for narrow-base transistors, fixing I_B and fixing V_{BE} would be equivalent in the absence of thermal effects. In practice, however, with fixed I_B, V_{BE} adjusts itself to cancel most of the thermal effect on the output resistance. This effect is easily demonstrated by comparing output characteristics measured with fixed I_B and with fixed V_{BE}. At high currents, the fixed-V_{BE} curves generally show much greater slopes.

4.3. Extraction Results

Measured thermal spreading resistance values were used with (4.1) and (4.6) to predict BJT small-signal behavior. Values for D_C and D_B were determined using SPICE. Figure 4.3 shows measured and simulated variation of normalized output resistance parameter $V_A' = r_o I_C - V_{CE}$ with I_C. The transistor had an emitter area of 23 x 23 $(\mu m)^2$, base width of 1.0 μm, and epitaxial-layer doping of 10^{15} cm^{-3}. The measured thermal spreading resistance R_{th} was 70 °C/W. The device was packaged in a 14-pin plastic DIP whose chip-to-ambient thermal

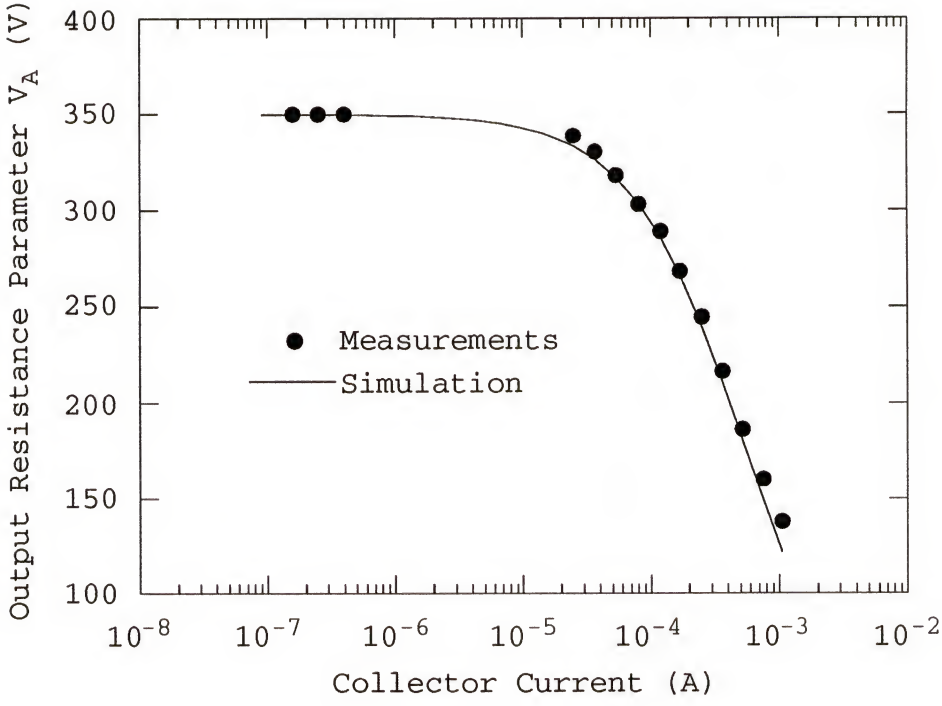


Fig. 4.3 Measured and simulated normalized output resistance parameter $V_A = I_C/g_{22} - V_{CE}$ for a 23 x 23 $(\mu m)^2$ BJT. R_{th} = 70 °C/W.

resistance was 110 °C/W. The measurements and the simulations were both arranged to include only the thermal spreading resistance. Figure 4.4 shows measured and simulated variation of g_{12} for the same transistor. In this case, an HP4145 Parameter Analyzer set to its maximum integration time was used to make the measurement. As discussed previously, this causes the package thermal impedance to contribute a component to the effective value of R_{th}. R_{th} was raised to 90 °C/W in the simulation to account for this effect.

The measurements and simulations in both figures agree well over many decades of current. The disagreement at high currents could be due to effects of parasitic emitter and base resistances or errors in the models for D_C and D_B at these current levels. Note that SPICE models predicts a constant value (V_{AF}) for normalized output resistance parameter and a zero or very small negative value for g_{12}. The much larger positive value for g_{12} seen here is important because it makes the input impedances depend strongly on loading condition at the opposite port.

Similar agreement between measured and simulated small-signal parameters have been seen with many other transistors. Table 4.1 shows R_{th} values found by the characterization methods described in this chapter. Of special interest is the high thermal spreading resistance seen with dielectrically isolated BJTs. Such transistors are isolated from the silicon

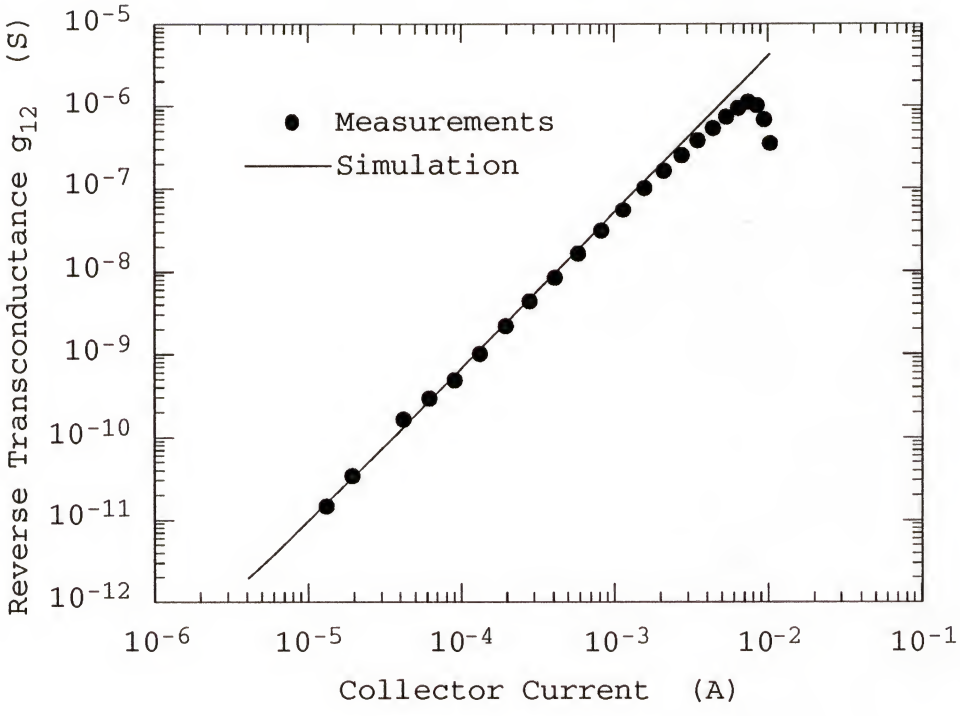


Fig. 4.4 Measured and simulated reverse transfer conductance g_{12} for 23 x 23 $(\mu m)^2$ BJT. Simulation used R_{th} = 90 °C/W to account for package thermal impedance affecting the measurement.

Table 4.1 Measured thermal spreading resistance for several transistor geometries.

Type	W (μm)	L (μm)	Number of emitter	Measured R_{th}
	9	9	1	118±10
	9	9	2x4	50
	54	9	2	50
	9	9	8	42
a	20	20	1	74±5
	20	20	2	57
	80	20	1	50±5
	20	20	4	43
	80	20	2	40±2
	20	20	8	31
	4	2	1	480±20
b	8	2	1	390±5
	10	2	1	350±20
	10	4	1	290±10
c	4	2	2	560±20
	8	2	1	435±20
d	10	6	1	280±20
e	22	11	1	327±3
	22	22	2	280±5
f	22	22	2	350±20

a: NPN, Base-width = 1 μm, N_{EPI} = $1x10^{15}$ cm^{-3}.

b: NPN, Base-width = 0.5 μm, N_{EPI} = $7x10^{15}$ cm^{-3}.

c: NPN, Base-width = 0.15 μm, N_{EPI} = $2x10^{16}$ cm^{-3}.

d: NPN, similar vertical dimension as c.

e: NPN, DI, Base-width = 0.8 μm, N_{EPI} = $1x10^{15}$ cm^{-3}.

f: PNP, DI, Base-width = 0.8 μm, N_{EPI} = $1x10^{16}$ cm^{-3}.

- W and L represent the dimensions of the mask opening.

substrate by a layer of silicon dioxide, whose thermal conductivity is only 1 % that of silicon. The R_{th} of a 23 x 23 $(\mu m)^2$ dielectrically isolated transistor was found to be 280 °C/W, four times that of the similar-size junction-isolated transistor previously described. Note that the thermal resistance of the vertical PNP transistor (dielectrically isolated) shows a comparable but slightly higher value of R_{th} compared to the same-dimension NPN transistor. Slightly higher value of R_{th} for the PNP transistor is consistent with the model described in Chapter 3 (Fig. 3.5) considering the higher epi-doping concentration.

4.4. Conclusion

Modeling of thermal spreading impedance is expected to become increasingly important in circuit and device simulation as transistor geometries are scaled down and current densities rise. This chapter presents practical ways to extract thermal spreading resistance of BJTs separately from package thermal impedance. These techniques will be especially useful for modeling of devices whose unusual geometries or structures (such as dielectric isolation) make analytical modeling of thermal impedance impractical.

CHAPTER 5
THERMAL EFFECTS IN BJT INTEGRATED CIRCUITs

5.1. Introduction

Circuit and device integration levels continue to improve with time. The resulting smaller device sizes result in higher thermal spreading resistances and the demand for higher performance gives rise to higher power densities in these devices. A result of these trends is towards higher junction temperature elevations in the devices. In addition, in new isolation technologies (e.g., trench isolation, dielectric isolation, and SOI), the poor thermal conductivity of silicon dioxide makes the temperature elevations far more significant.

As discussed in the previous chapters, for modest powers, even adjacent devices are thermally decoupled from each other [Fo91a] so that localized heating becomes intrinsic to the BJT itself. This work focuses on this intrinsic self-heating effects which can not be avoided through changing the layout of the IC.

It was shown in [De89] that neglecting intrinsic local heating effects caused errors up to 4 % in the predicted

delay of a simple BJT logic gate. Local heating causes much larger errors in some widely used high accuracy monolithic linear integrated circuits such as D/A converters, instrumentation amplifiers, analog multipliers, logarithmic amplifiers, current mirrors, and precision voltage references. Local heating can also cause errors in op-amp practice as it can lower the low-frequency open-loop gain, cause distortion, and change the phase response. Therefore, it is important for designers to have good insight into the circuit implications of local-heating effects, as even well-executed designs can fall prey to these effects [To90].

Knowledge of the implications of thermal effects in integrated circuits not only provides the design principles to reduce thermal effects in IC design but can be used to avoid them for better IC design. The object of this chapter is to provide general insight into local-heating effects on the DC and small-signal performance of bipolar integrated circuits and to identify the circuit conditions under which self-heating causes significant errors in simulation. The thermal impedance model described in Chapter 3 has been implemented in SPICE DC and AC analysis by D. Zweidinger [Zw92]. This improved BJT model is used to verify the discussions in this chapter.

In Section 5.2, the improved SPICE BJT model [Fo91c] is introduced which includes a brief description of how the

thermal models are incorporated in the DC and AC routines, and discusses the three levels of thermal modeling available to the user [Zw92]. In addition, a series of simulations are compared to measurements to demonstrate the improved accuracy afforded by the thermally enhanced BJT model.

In Section 5.3, DC thermal effects are considered in two aspects, magnification of the increment in collector current due to the mismatch in collector-base voltage and collector current deviation from near-exact exponential dependence on base-emitter voltage. For small-signals, the effects of self-heating on the input and output resistances and the effects of series emitter, base, and collector resistances are discussed. The effects of discrepancies in the input and output impedance on other small-signal behavior are also discussed. Simulation results of a few circuit configurations, based on modified and ordinary SPICE, are provided for verification and demonstration purposes. Design strategies to avoid significant thermal problems are also discussed in this section.

In Section 5.4, the work discussed in this chapter are summarized.

5.2. Modified SPICE BJT model

From Chapter 4, a simple model for the thermal impedance of any rectangular emitter geometry is given by

$$Z_{th} = R_{th}\exp\left(-r_{eff}\sqrt{s/k}\right) \tag{5.1a}$$

where

$$r_{eff} = \frac{1}{2\pi K R_{th}} \tag{5.1b}$$

and R_{th} is given as

$$R_{th} = 1/4\,\pi K\sqrt{WL}\,f_1\,f_2 \tag{5.1c}$$

where

$$f_1(d,h) = (0.058d+0.14)h+0.34d+0.28, \tag{5.1d}$$

$$f_2(a) = 0.98+0.043a-(6.9\cdot10^{-4})a^2+(3.9\cdot10^{-6})a^3, \tag{5.1e}$$

$$a = W/L, \tag{5.1f}$$

$$d = D/\sqrt{WL}, \tag{5.1g}$$

$$h = \frac{1}{\sqrt{WL}}\sqrt{\frac{2\varepsilon(V_o + V_{CB})}{qN_{EPI}}}. \tag{5.1h}$$

V_o is the collector-base junction built-in potential. This model, which is valid for junction-isolated devices, has been implemented in SPICE DC and AC analysis. MicroSim's PSpice [Ps89] was used as the initial code for the modifications, using the device models option. The SPICE DC analyses were modified in three respects to include thermal effects. First, in the operating-point routine, the large-signal currents and all temperature-dependent parameters were adjusted to reflect

a raised local temperature for each device using the standard SPICE equations. Second, the admittances were modified to include thermal effects using (3.6). Finally, the increase in temperature from one iteration to the next was limited to improve convergence of the operating point and prevent thermal runaway. Figure 5.1 shows a flowchart of the modified subroutine to load the conductance matrix, which is the heart of the DC analysis and operating-point computation. The modifications to the AC analysis routine consist primarily of changes to the admittances, which are similar to those in the DC analysis.

The modified SPICE requires a few parameters to be added to the BJT model description. Three levels of thermal modeling are provided to the user. Each level requires a different way of specifying thermal impedance parameters in the model specification. The first level is fully predictive and requires only information about the device geometry and doping. From (5.1), the thermal impedance can be specified for a given W, L, D, and N_{EPI}. This predictive approach allows circuits to be simulated concurrently with a developing process.

The second method of determining Z_{th} uses a value of R_{th} extracted from measurements. The frequency response is then predicted using (5.1a) and (5.1b). Suggested techniques to measure R_{th} are given in Chapter 4. This modeling approach can

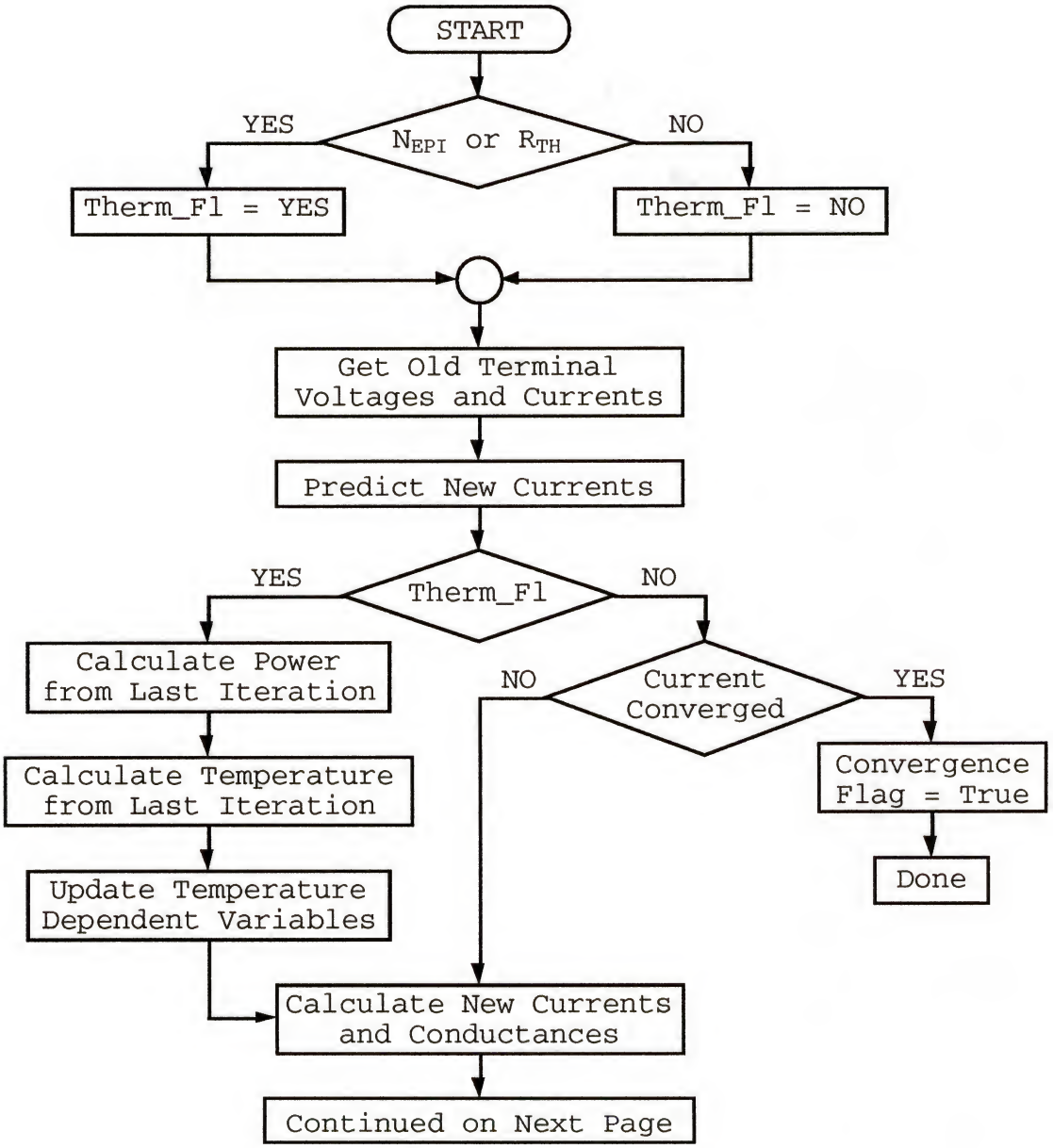


Fig. 5.1. Flowchart of BJT DC analysis.

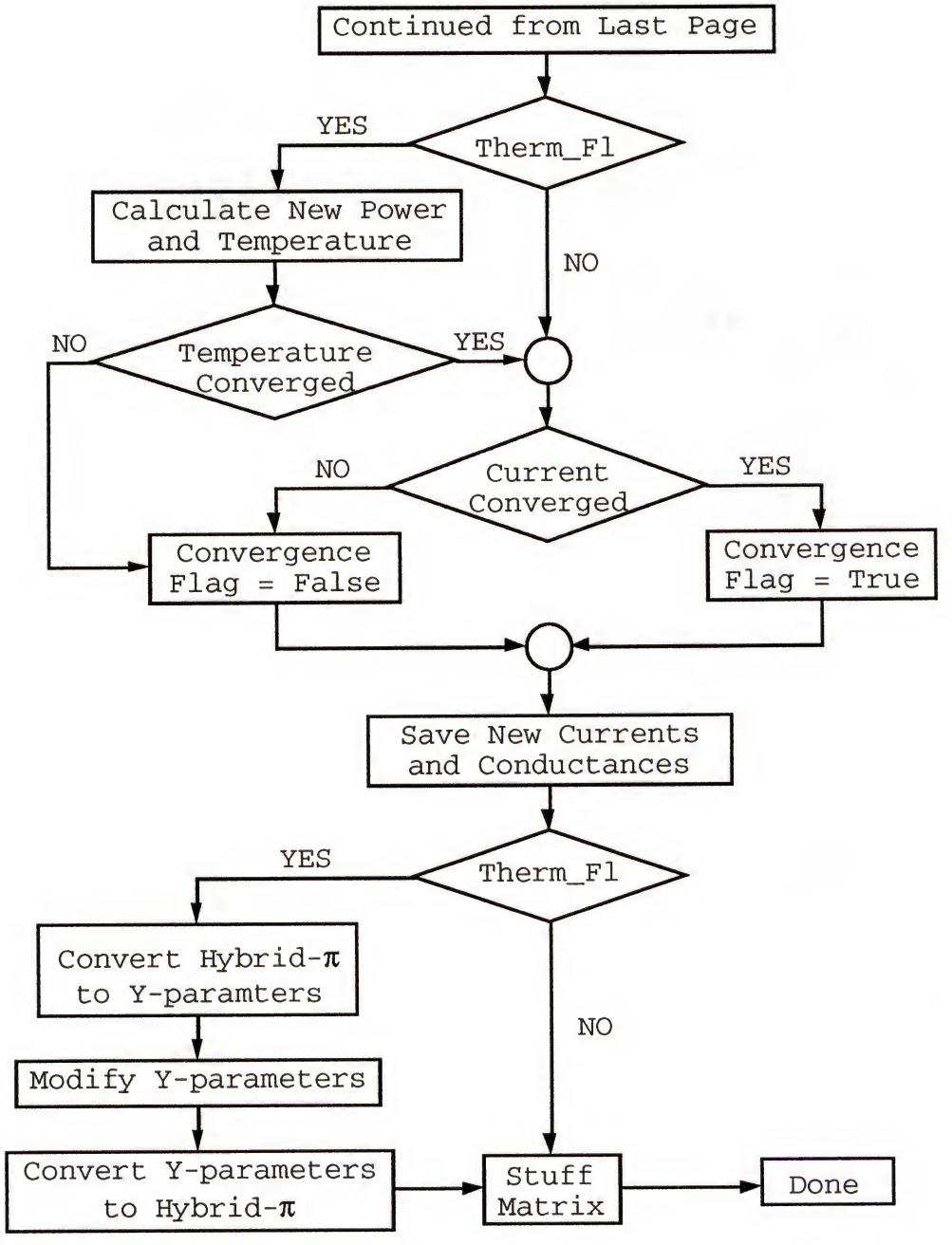


Fig. 5.1. continued.

give greater accuracy than the predictive approach when good measurements of R_{th} can be made.

The third approach uses a single-pole model, in which Z_{th} is specified by providing the real part, R_{th}, and the imaginary part, C_{th}. When specified in the manner, simulations are only accurate over a limited frequency range.

A series of simulations of I_C vs. V_{CE}, Early voltage V_A vs. I_C, g_{12} vs. I_C, and Y_{22} vs. frequency are compared to measurements in Figs. 5.2 - 5.5, respectively. The transistor used was rather big. Its thermal spreading resistance R_{th} was near 70 °C/W. The device had emitter length and width of 23 μm, collector-base junction depth of 3.0 μm, and epi-doping N_{EPI} of 10^{15} cm^{-3}. The fully predictive model was used for the simulations. In Fig. 5.2, as expected, measurements and modified SPICE simulations predict higher collector current and much higher I_C - V_{CE} slopes than ordinary SPICE does. The increase in I_C - V_{CE} slope for high I_C can be seen more clearly in Fig. 5.3. The Early voltage V_A, which is predicted to be constant based on the electrical-only model, approaches zero for high I_C as the thermally induced variation dominates the I_C - V_{CE} slope. Note that this transistor has a rather high electrical-only Early voltage V_A so that even though R_{th} is only 70 °C/W, the output resistance drops to half of what the electrical-only model predicts at only $I_C \cong 400$ μA.

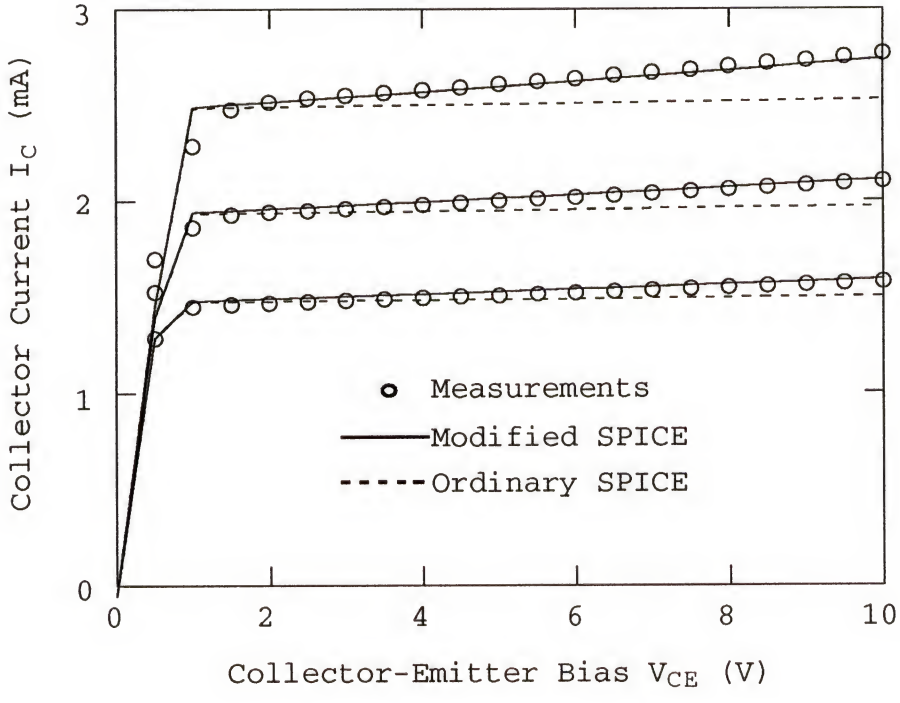


Fig. 5.2. Typical measured and simulated collector currents as a function of collector-emitter voltage.

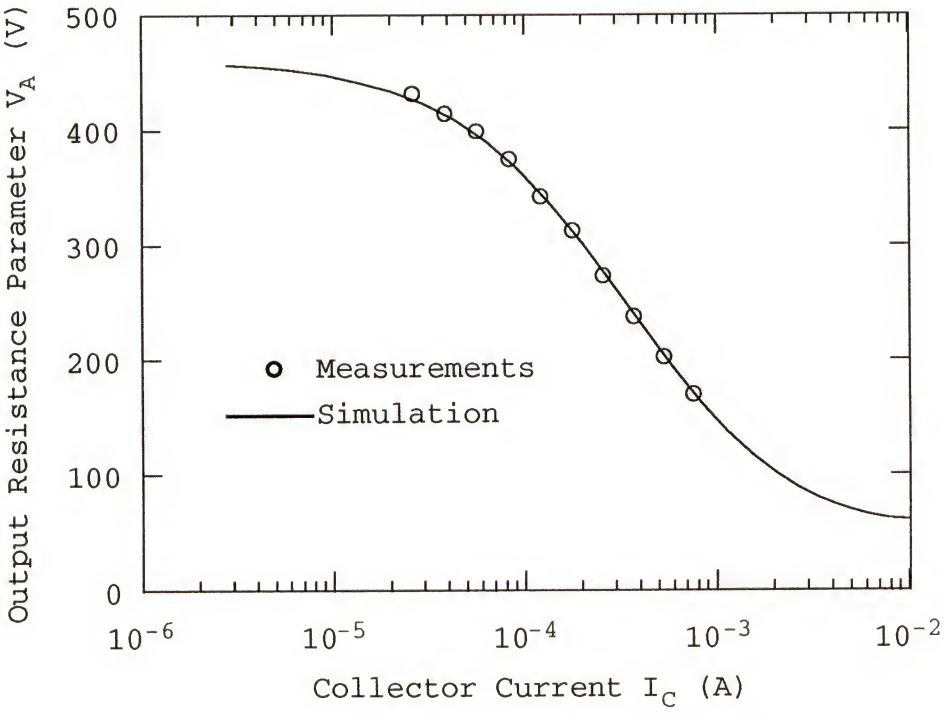


Fig. 5.3. Normalized output resistance $V_A = I_C/g_{22} - V_{CE}$ as a function of I_C, measured and simulated. V_{CE} was set to 6V. Note that unmodified Gummel-Poon model predicts a constant for this parameter.

In Fig. 5.4, measurement and modified SPICE simulation of g_{12} as a function of I_C show good agreement up to reasonably high currents. The discrepancies at high currents are probably caused by the inaccuracies in the model for the temperature coefficient of I_B and by inexact parameters for base and emitter parasitic resistances that are critical at those currents. Note that g_{12} is considered so small that in ordinary SPICE this parameter is not printed in the operating point output. The components of base current which depend on collector emitter voltage are the recombination in quasi-neutral base (QNB) and the generation in collector junction space-charge region (SCR) [Sz81]. These current components are typically negligible for contemporary vertical BJTs, and tend to reduce base current with increasing V_{CE}. In principle, modulation in QNB-width modulates the base-spreading resistance also, so that base current decreases with increasing V_{CE}. Therefore, g_{12} predicted by electrical-only model is negative and small. Impact ionization can affect the base-current significantly [Po89], especially for advanced devices, but again, drives g_{12} to negative value.

Figure 5.5 shows measured and simulated value of Y_{22}, both magnitude and phase for I_C = 1.5 mA and V_{CE} = 6 V. The discrepancies in the phase of Y_{22} at high frequencies (~ 10^7 Hz) are probably the measurement errors (e.g., components become more inductive). Figure 5.5a shows that, in this

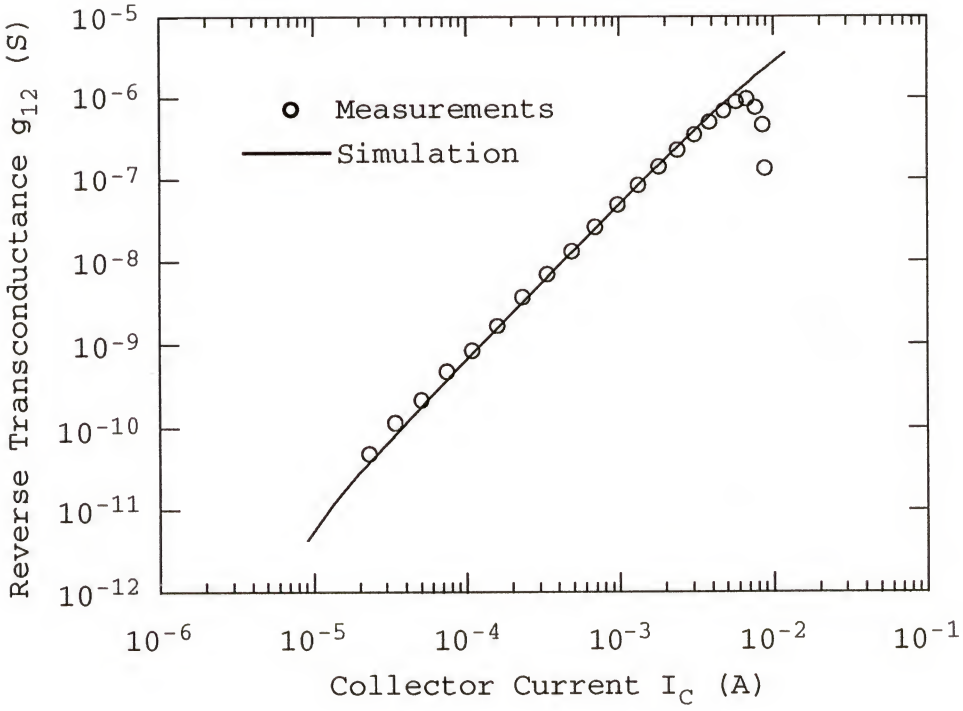
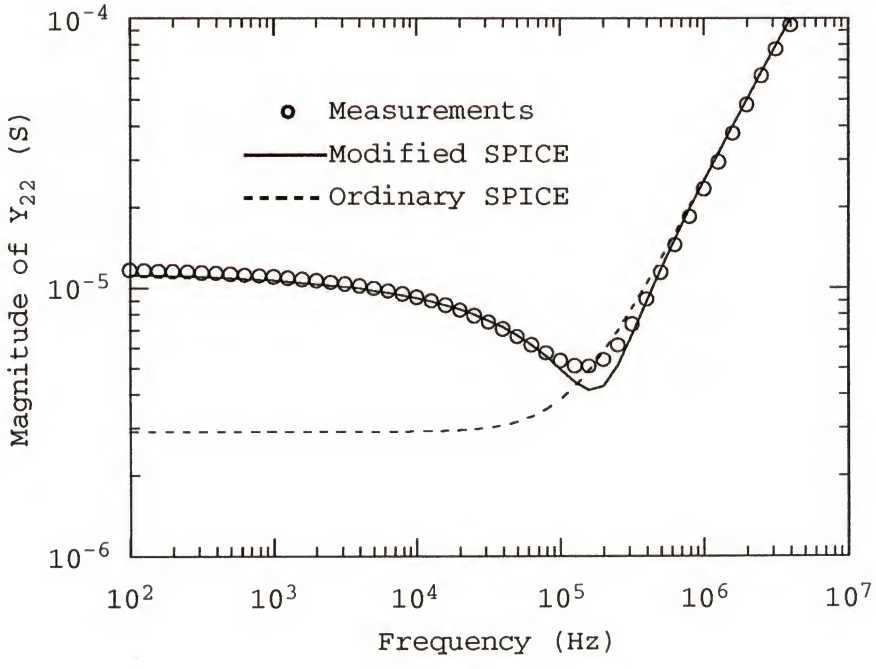
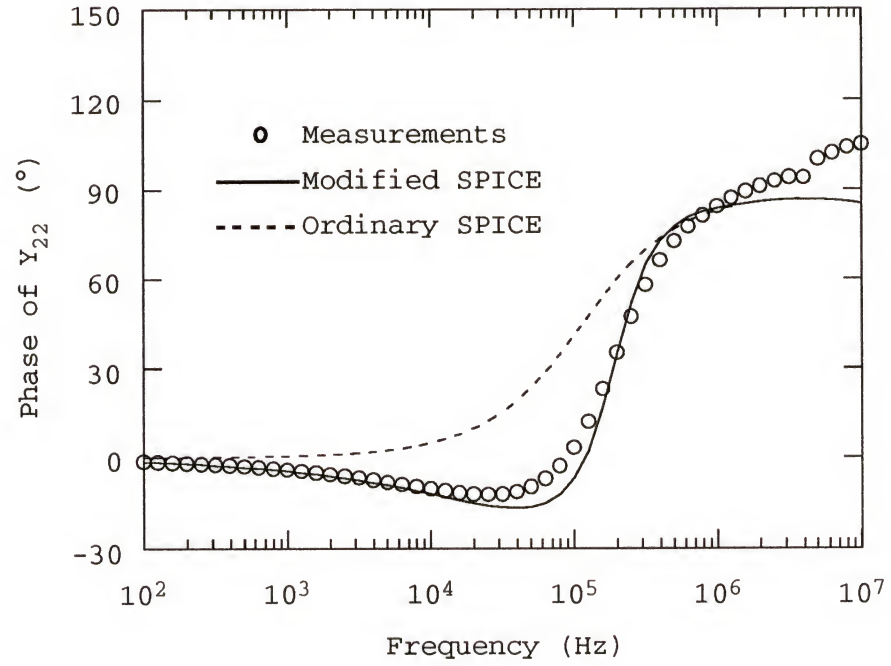


Fig. 5.4. Common-emitter reverse transconductance as a function of collector current, measured and simulated. Unmodified SPICE gives a small negative value for this parameter.



(a)



(b)

Fig. 5.5. Measured and simulated output admittance. (a) Magnitude. (b) Phase. V_{CE} = 6 V and I_C = 1.5 mA.

transistor, thermal effects disappear for frequencies higher than about 300 kHz.

As Figs. 5.2 - 5.5 show the thermally modified SPICE simulations agree well with measurements, demonstrating the superiority of the new model compared to the electrical-only model. Since the measurements and simulations show good agreements on the device level, in the following section, circuit simulation results predicted by the modified SPICE will be assumed to be the same as measurements.

5.3. Thermal effects in integrated circuits

The large-signal effects of self-heating on IC performance will be considered first. Then the effects on small-signal performances will be considered. Circuit simulations will also be included for verification and demonstration. The NPN transistors used for the simulations had the following specifications: emitter length of 4 μm, width of 10 μm, collector-base junction depth of 0.4 μm, and $N_{EPI} = 7 \times 10^{15}$ cm^{-3}. The value V_A = 150 V was extracted at low currents. This device had about the same current-driving capabilities as the 23x23 $(\mu m)^2$ transistor used in Section 5.2 as this device was more heavily doped than the other. However, the thermal resistance of this transistor was 290 °C/W, higher than the 70 °C/W, R_{th} of the other device. Therefore, more significant thermal effects are expected in

circuits using this device than the one used in Section 5.2.

The thermal impedance model presented in Chapter 3 is not directly applicable for lateral PNP transistors. However, it is valid for vertical PNP transistors. As discussed in Chapter 4, vertical PNP transistors have been shown to exhibit similar thermal spreading resistances as NPN for a given emitter area. In the simulations to follow, the model parameters of the 4x10 $(\mu m)^2$ NPN transistor are also used for the PNP transistors. Due to lower hole mobility, actual PNP transistors would show lower collector current and larger transit time τ_F than the same dimension NPN for a given base-emitter voltage. However, using the same parameters makes little difference on the simulation results.

5.3.1. DC thermal effects in IC

The increment in emitter-junction temperature over the ambient due to collector power dissipation is given by $\Delta T \cong R_{th} I_C V_{CE}$ so that the DC thermal effects are increasingly significant toward higher R_{th}, I_C, and V_{CE}. Typically, the terminal currents increase roughly 7%/°C when the base-emitter voltage is fixed, and roughly 0.5 %/°C when the base current is fixed [Sz81]. For the current-driven case, the base-emitter voltage varies roughly -2 mV/°C. More exactly, the temperature coefficient of terminal currents and base-emitter voltage is a decreasing linear function of V_{BE}, given

by (4.4), (4.8), and (14) in [Sl76]. Therefore, thermal effects are more significant at low V_{BE} for the same amount of power dissipated.

With regard to circuit design, the problems caused by the DC thermal effects can generally be classified into two categories. First, thermal effects magnify the dependence of collector current on V_{CE}, and second, thermal effects make the terminal currents depart from the exact exponential dependence on the base-emitter voltage.

The significance of the former effect, which is most significant for the voltage-driven case, will first be considered through the example of a current mirror. The current mirror is an ubiquitous building block in analog integrated circuits. Figure 5.6 shows a basic current-mirror circuit, which is susceptible to an offset-current (I_{out} - I_{ref}) error not only due to the limited Early voltage, but also due to local thermal feedback.

Assuming the transistors are perfectly matched, an offset current can be developed due to unequal voltages on the collectors of Q_1 and Q_2. An electrical-only BJT model would predict an offset current error resulting from the difference in base-width modulation while the actual value of offset current will also be affected by thermal effects. For I_{ref} = 1 mA and V_{C2} - V_{C1} = 10 V, the actual temperature of Q_2 is higher than Q_1 by roughly $R_{th}I_{ref}(V_{C2} - V_{C1})$ which is about

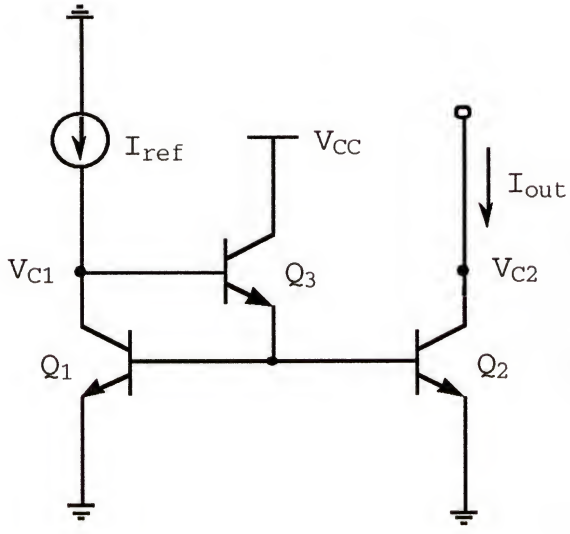


Fig. 5.6. Circuit diagram of a simple current mirror.

2.9 °C. This is effectively the same condition as the base-emitter voltage of Q_2 is driven by roughly 5.8 mV (V_{BE} varies by about -2 mV/°C) harder than Q_1. Figure 5.7 shows the offset current (I_{out} - I_{ref}) as a function of I_{ref} for a differential collector voltage (V_{C2} - V_{C1}) of 10 V based on modified and ordinary SPICE simulations. As can be seen in Fig. 5.7, the errors between the two offset currents become increasingly significant toward higher differential power. At I_C = 1 mA, ordinary SPICE predicts an offset current of 75 μA while the actual value is 226 μA, a ratio of about 1 to 3.

In above example, one might suggest that the error can be avoided by extracting V_A at the actual operating point of those transistors. However, the extracted Early voltage varies strongly as a function of I_C due to thermal effects. Early voltage can even change significantly over the collector-emitter voltage swing. Therefore, extracting V_A at the actual operating current is not practical for the purpose of predicting circuit performance.

In fact, thermal modeling is necessary to avoid errors even for the simpliest case. Fig. 5.8 show a measured I_C versus V_{CE} of the 4x10 (μm) NPN transistor, for the cases base is driven by a voltage source and current source. As can be seen in the figure, collector currents which are same at low V_{CE} diverge significantly toward high V_{CE}. These diversions are due to the discrepancies in temperature coefficient of

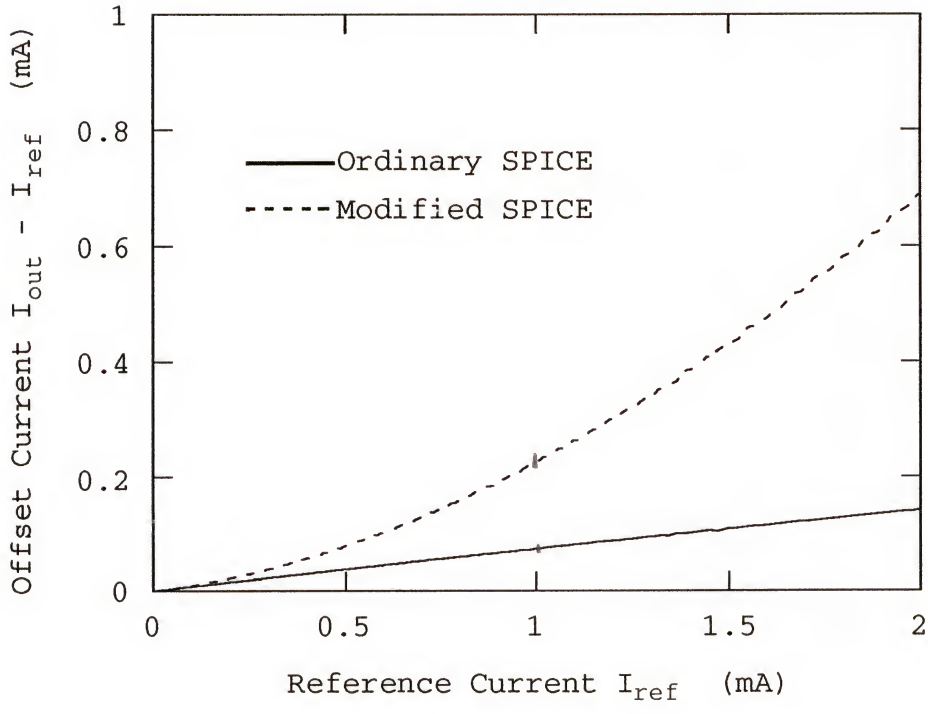


Fig. 5.7. Offset current (I_{out} - I_{ref}) versus mirror current (I_{ref}) of the simple current mirror for differential collector voltages (V_{C2} - V_{C1}) of 10 V.

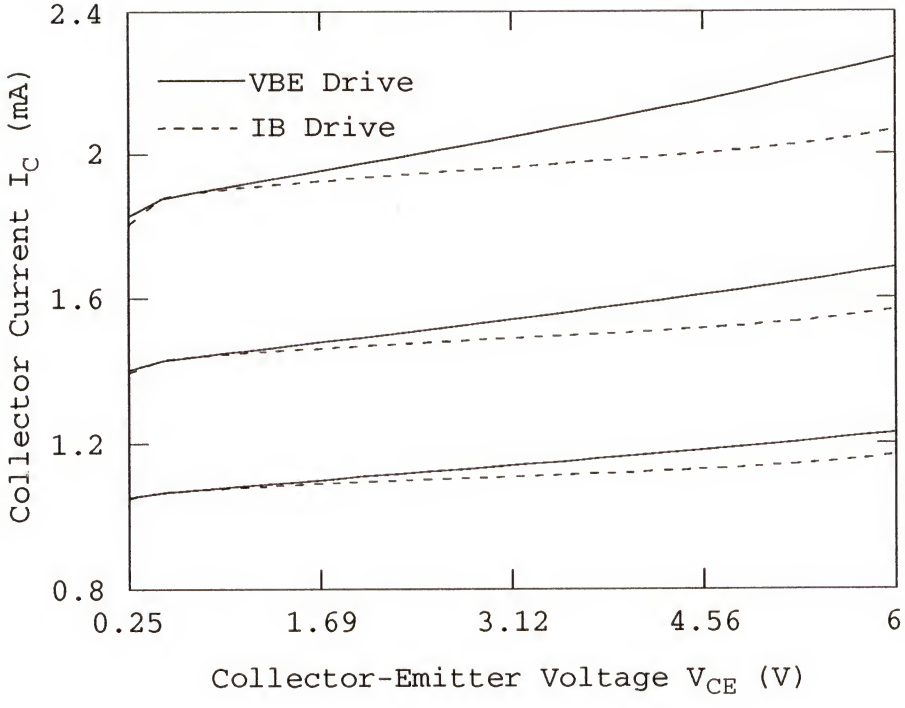


Fig. 5.8 Typical measured collector currents as a function of collector-emitter voltage. Emitter area = 4x10 $(\mu m)^2$ and N_{EPI} = $7x10^{15}$ cm^{-3}.

collector current which depend on the bias conditions. Therefore, even at a given collector current I_C, a single value V_A can not satisfy the output characteristic as the small-signal slope of the output depends on the bias condtion.

Fig. 5.9 shows a self-biasing V_{BE}-based current reference which is another circuit susceptible to the Early voltage degradation. In Fig. 5.9, changing the supply voltage, V_{CC}, mostly varies the collector-base voltage of transistor Q_3, V_{CE3}, and thermal effects make the current mirror output I_{C3} to be strongly dependent on V_{CE3}. For V_{CC} = 12 V, ordinary SPICE predicts 0.09 (0.003) for the sensitivity of the output current I_{out} to the supply voltage V_{CC} which is defined as $(V_{CC}/I_{out})(\partial I_{out}/\partial V_{CC})$. For the same circuit, modified SPICE predicts 0.45 (0.021), which is more than five times higher than what the ordinary SPICE predicts. This amount of error can turn a useful circuit into a totally inadequate one.

In both of above examples, increasing the base or emitter resistance can help reduce the errors. The negative electrical feedback across R_B or R_E opposes the increase in terminal currents. This will be explained more detail in the section on small-signal analysis. In the supply-independent current source, replacing the current mirror pair Q_3 and Q_4

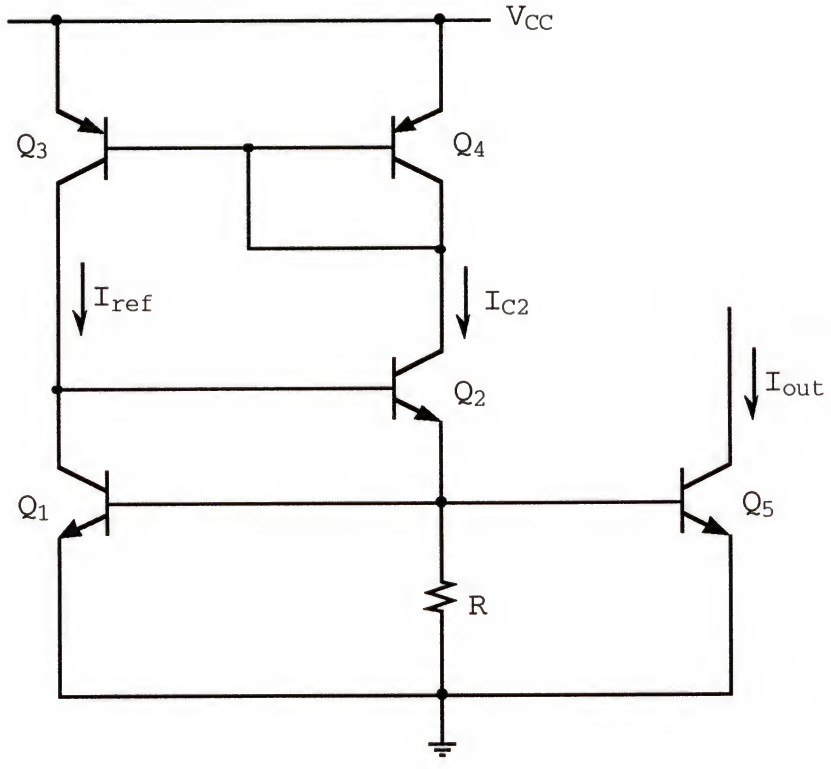


Fig. 5.9. Circuit diagram of a self-biasing V_{BE} reference.

with a cascode or Wilson current source can greatly reduce the sensitivities. In a cascode source, such as in Fig. 5.10, the output current is mostly determined by the common-emitter transistor Q_3. Changing the output voltage V_{C2} primarily changes the power in common-base transistor Q_2. Since Q_2 is driven with nearly constant current, the self-heating of Q_2 mostly changes its base-emitter voltage, and has little effect on the output current.

A class of circuits susceptible to the deviations from the exponential dependence of collector current to the base-emitter voltage are the translinear circuits. Translinear circuits exploit the near-exact proportionality of the transconductance of a BJT to its collector current (or the near-exponential dependence of the collector current on base-emitter voltage) hence the term translinear.

Figure 5.11 shows a four-quadrant multiplier cell. Depending on the input voltages v_1 and v_2, collector currents will vary between 0 and 2 mA. Therefore, the temperature of each transistor can vary substantially as a function of inputs. The variations in the temperature of each device introduce extra non-proportionality of transconductance to its collector currents so that the differential output current $(I_1 - I_2)$ deviates more from linear behavior. Figure 5.12 compares the deviation of differential output current $(I_1 - I_2)$ from the linear behavior as a function of input

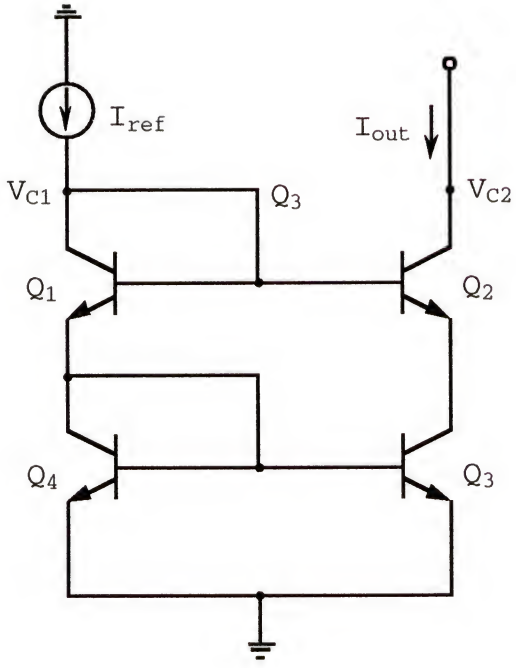


Fig. 5.10 Circuit diagram of a cascode current source.

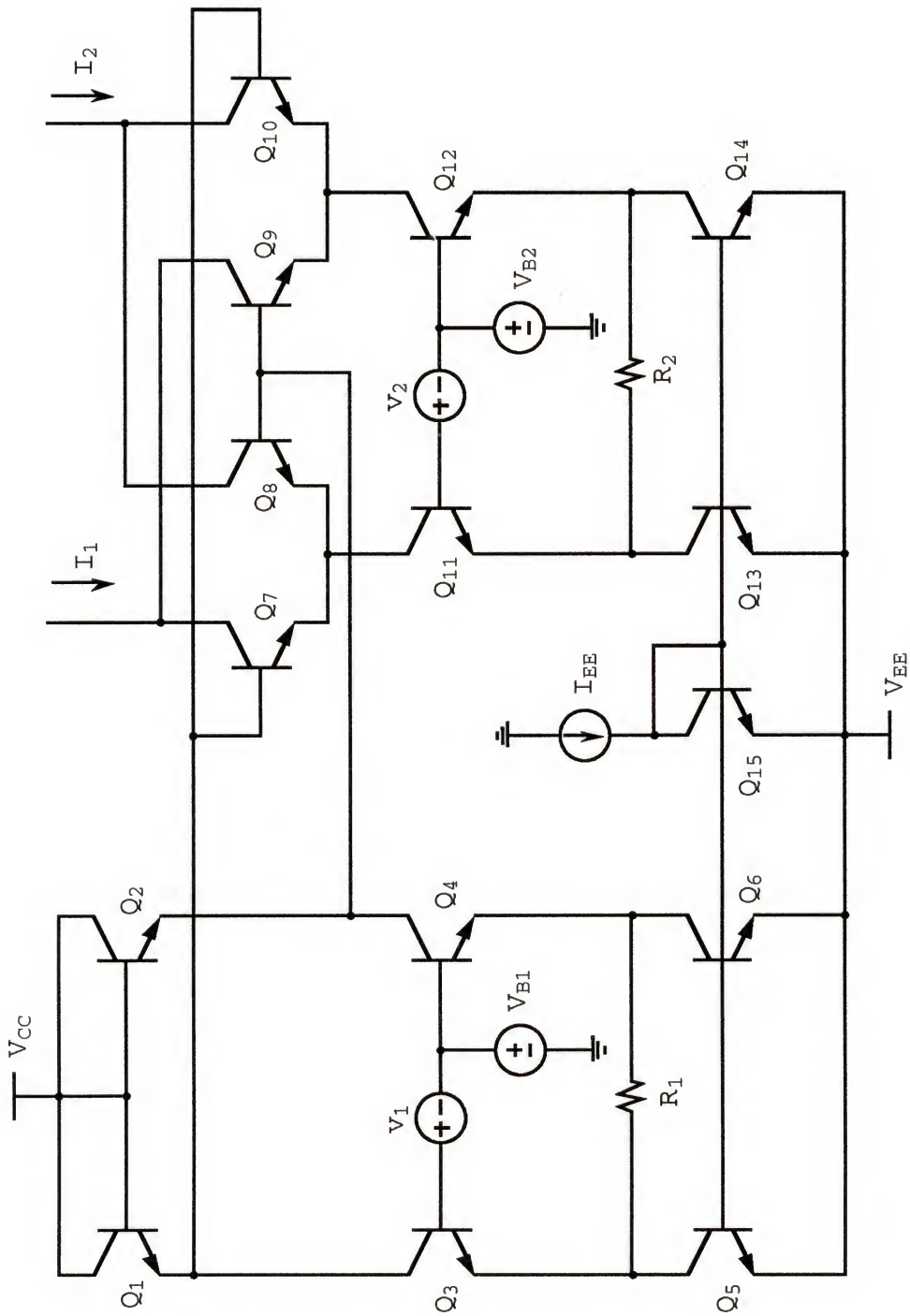


Fig. 5.11 Circuit diagram of a four quadrant multiplier.

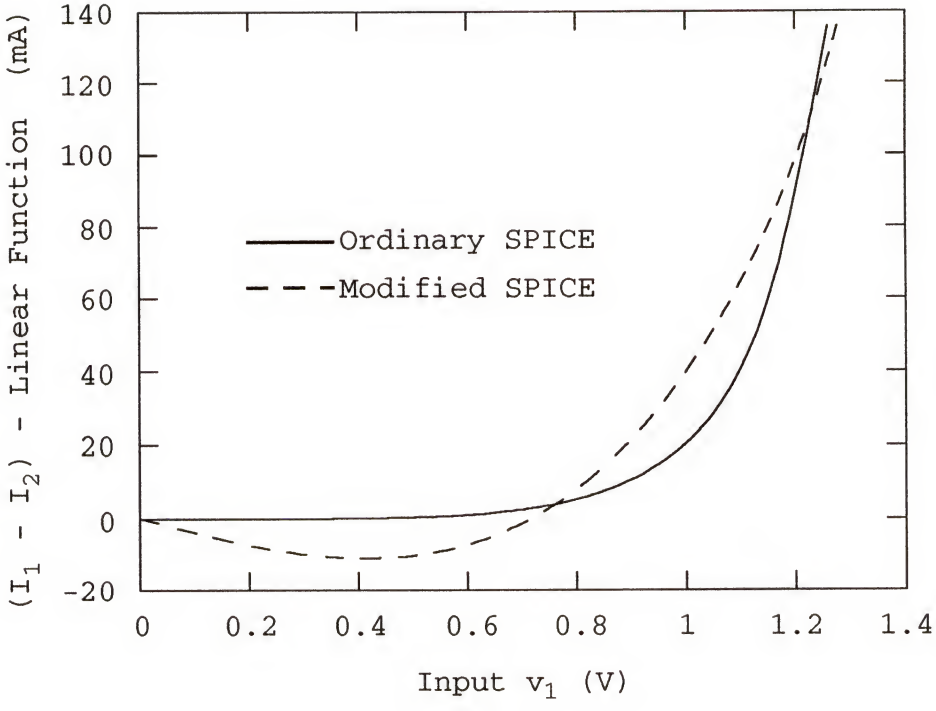


Fig. 5.12. Deviation of the differential output current (I_1 - I_2) from the linear behavior as a function of input v_1 for v_2 = 0.5 V.

voltage v_1 for v_2 = 0.5 V, based on ordinary and modified SPICE simulations. As can be seen in Fig. 5.12, thermal effects make the differential output current a nonlinear function of the inputs in the region where ordinary SPICE predicts a linear behavior. This is direct indication of a large amount of distortion.

Figure 5.13 shows a typical band-gap voltage reference circuit. In the figure, the emitter area of Q_1 is ten times that of Q_2. Resistor R_2 is made to be $10 \cdot R_3$, so that $I_{C2} = 10 \cdot I_{C1}$. Note that the ten times smaller transistor Q_2 dissipates ten times more power than Q_1. If the two transistors Q_1 and Q_2 are assumed to operate at the same temperature, the output voltage V_{ref} can be expressed as

$$V_{ref} = V_{BE1} + V_T \left[\ln\left(\frac{I_{E2}}{I_{E1}} \frac{I_{S1}}{I_{S2}}\right)\left(1 + \frac{R_2}{R_1}\right)\right]. \qquad (5.2)$$

The value of V_{ref} is significant, as it is the point where the negative and positive temperature coefficient of the first and second term in (5.2) cancels out, so that the temperature coefficient of reference voltage becomes zero at V_{ref}. This circuit is designed to have zero temperature coefficient at temperature T = 0 °C with V_{ref} = 1.293 V. Figure 5.14 shows the simulation results based on ordinary (dashed line) and modified (solid line) SPICE for $I_{C2} \cong 1$ mA and V_{CC} = 5 V. Thermal effects lead to about 10 mV discrepancies in the

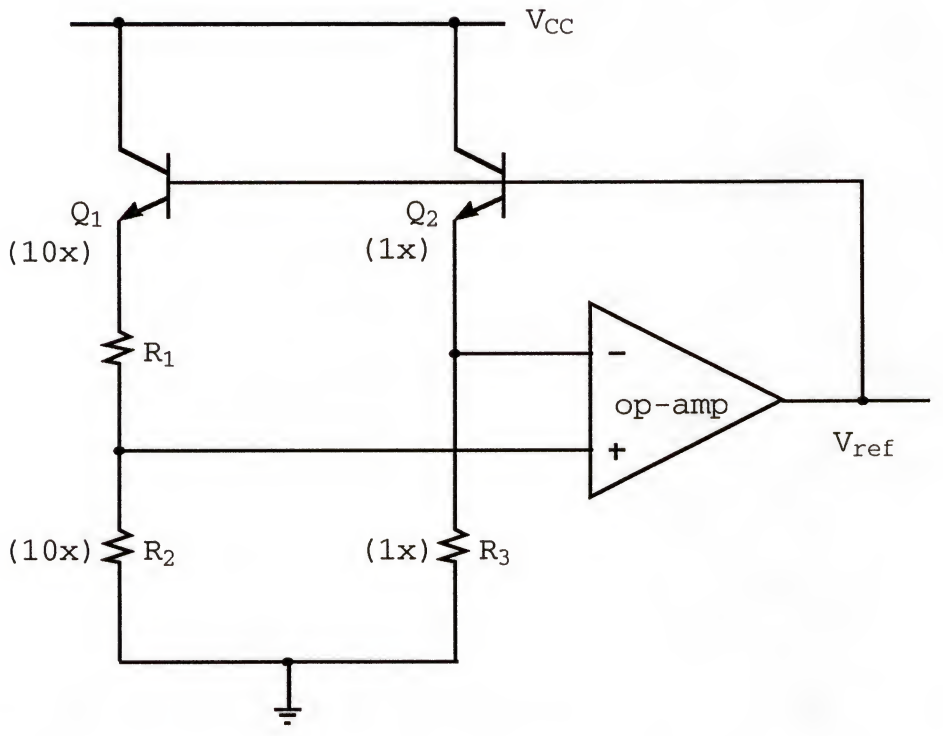


Fig. 5.13. Circuit diagram of a band-gap voltage reference.

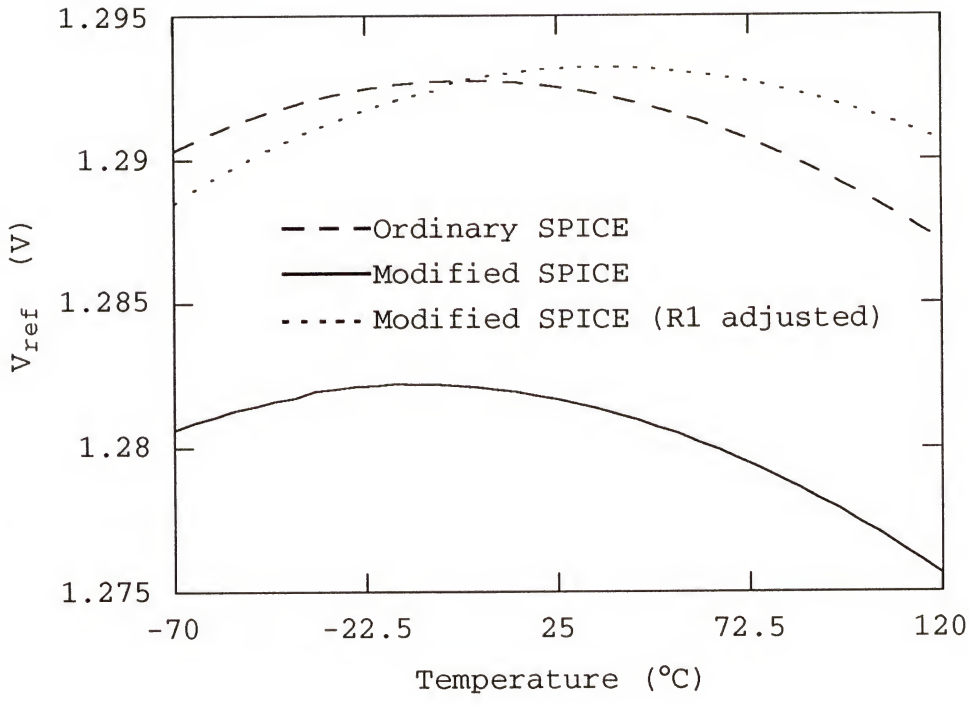


Fig. 5.14. Output voltage of band-gap voltage reference as a function of temperature.

output voltage V_{ref} and 5 °C displacement in T_o, which is defined as the temperature at which $\partial V_{ref}/\partial T = 0$. These errors are caused by the deviation of I_C from the near-exponential V_{BE} dependence. A popular way to tune the temperature coefficient of the band-gap reference is to trim the output to the value of V_{ref} [Ku73, Br74], assuming the right output value will give the right temperature coefficient. Here, resistor R_1 is adjusted to match the output voltage at T = 0 °C to the value predicted by ordinary SPICE (1.293V). Figure 5.14 (dotted line) shows the result of modified SPICE simulation after R_1 has been adjusted. Corrections without understanding the error source can worsen the errors. As can be seen in Fig. 5.12, this procedure actually displaces T_o by more than 30 °C.

In this section, circuit implication of DC thermal effects have been studied. The behavior of device level thermal effect that can give significant errors in many analog circuits are the strong dependence of I_C on V_{CE} and I_C being a less-exact exponential function of V_{BE}. These behaviors lead to errors in many precision analog circuits such as increasing the offset currents, degrading the supply voltage sensitivity, distorting the output behavior, and degrading the temperature dependence of output voltage. In general, large-signal thermal effects are insignificant if the power dissipation is small. The errors caused by stronger

dependence of I_C on V_{CE} can be reduced considerably through electrical negative feedback or cascoding.

5.3.2. Small-signal thermal effects in IC

In small-signal analysis, the thermally induced variations in terminal currents or voltages lead to change in the y-parameters. The common-emitter y-parameters corrected for thermal feedback, (3.6), are restated here

$$Y_{mn} = \frac{Y_{mnE} + D_m Z_{th} I_m I_n}{1 - D_m Z_{th} P} \qquad (5.3)$$

where m or n = 1 for the base or 2 for the collector, Y_{mnE} is the uncorrected electrical Y parameter, Z_{th} is the thermal impedance, and P is the dissipated power. D_2 is the fractional temperature coefficient of the collector current, and D_1 is that of the base current. Typically, at low frequencies Y_{12E} and Y_{22E} are both small so that the thermal term, $D_m Z_{th} I_m I_n$, can easily dominate. Conversely, Y_{11E} and Y_{21E} are larger to start with, so that thermal effects on these parameters are significant only for high powers in which case the denominators become considerably less than one. For example, for the 4x10 (μm^2) transistor used here, the denominator of (5.3) becomes 10% less than 1 at $I_C \cong 1.5$ mA assuming $D_2 \cong 4.5$ %/°C and $V_{CE} = 5$ V. At $I_C = 1.5$ mA, the

second term of (5.3) is only about $6x10^{-4}$% of the first term for both g_{11} and g_{21}.

The denominator of (5.3) is actually a measure of the significance of DC thermal effects. $D_mR_{th}P$ term includes all the parameters determining the DC thermal effects. If $D_mR_{th}P$ approaches 1, the transistor will get into thermal runaway. Thermal runaway requires very high powers and typical integrated circuits are unlikely to operate in that region, unless the thermal impedances are extremely high. Thermal runaway is an important aspect of thermal effects, and has been well known in power electronics [Sc67, Ho74, Ar74, La81]. This work excludes such a high power local-heating effects. Therefore, for the modest power, assuming $(1-I_mR_{th}P) \cong 1$, the significance of small-signal thermal effects is determined by the size of R_{th}, I_B, I_C, and the temperature coefficient of I_B and I_C.

Thermal effects on the low-frequency small-signal behavior of integrated circuits, will be considered here first. Self-heating affects y-parameters the most at frequencies close to DC, where (5.3) simplifies to

$$g_{mn} \cong g_{mnE} + D_mR_{th}I_mI_n. \qquad (5.4)$$

As mentioned above, g_{12} and g_{22} are the DC small-signal parameters most affected by thermal feedback. As mentioned previously, any other physical mechanisms except thermal

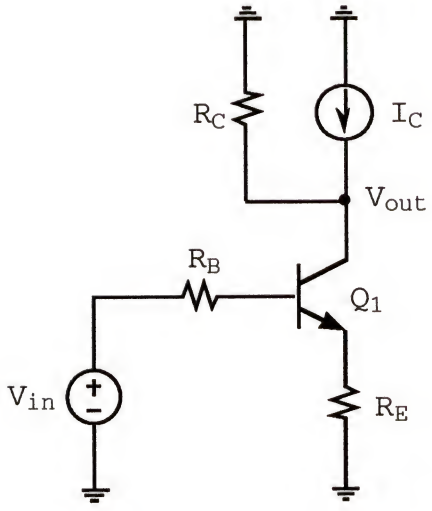
effects tend to make g_{12} negative. However, thermal effects make g_{12} positive even at low to medium currents since g_{12E} is so small in today's typical BJTs.

Figures 5.15a and 5.15b show a one-transistor common-emitter voltage-gain stage and its hybrid-π small-signal equivalent circuit. In Fig. 5.15b, $r_\pi = 1/g_{11}$, $g_\mu = g_{12}$, $g_m = g_{21}$, and $r_o = 1/g_{22}$, so that thermal effects can be incorporated into these conductances based on (5.4), and R_E, R_B, and R_C represent the total resistance looking out at emitter, base, and collector port, respectively, including parasitic resistances. From (5.4), r_o can be expressed as

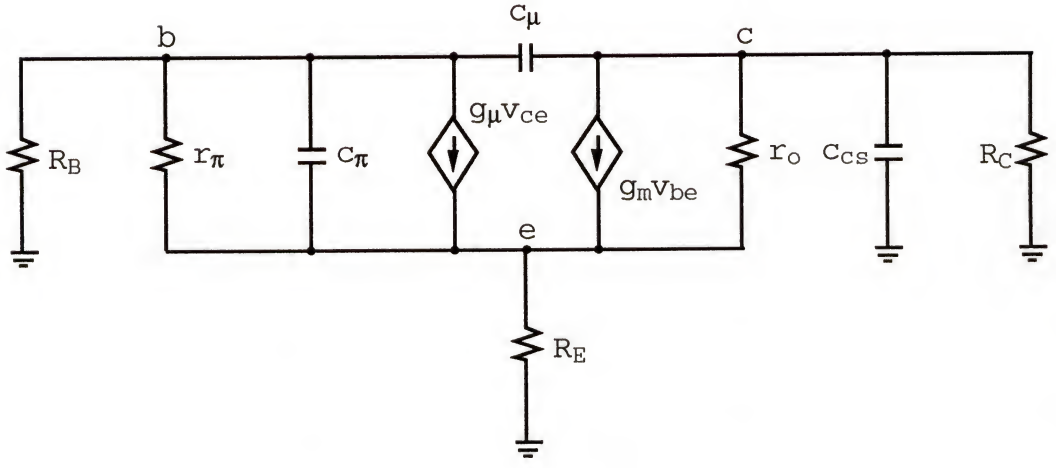
$$r_o = \frac{r_{oE}}{1 + D_2 R_{th} I_C V_A} \tag{5.5}$$

where r_{oE} is the electrical output resistance which is defined as V_A/I_C. Here V_A represents the Early voltage measured at low currents where the output conductance is controlled entirely by electrical effects. Note that the dimensionless term $D_2 R_{th} I_C V_A$ in (5.5) is a measure of the significance of thermal effects on r_o. This term involves all the factors determining the thermal degradation of r_o. The greater this term, the more r_o is degraded. r_o drops to half of what the electrical-only model predicts when $D_2 R_{th} I_C V_A = 1$.

From Fig. 5.15b, the output resistance at the collector is given by



(a)



(b)

Fig. 5.15. A common-emitter voltage gain stage. (a) Circuit diagram. (b) Equivalent circuit diagram.

$$R_{out} = r_o + R_E//R_B + \left[\frac{R_E + g_\mu r_o(R_E + R_B)}{r_\pi + (R_E + R_B)(1 - \beta_o g_\mu r_o)}\right]\left(\frac{r_\pi R_E}{R_E + R_B} + \beta_o r_o\right) \quad (5.6)$$

where β_o is assumed to be equal to $g_m r_\pi$. Based on (5.4), assuming g_{12E} is negligible, and assuming the small-signal current gain β_o equals the DC gain I_C/I_B, the $1-\beta_o r_o g_\mu$ term in (5.6) can be expressed as

$$1 - \beta_o r_o g_\mu \cong \frac{1 + D\beta V_A R_{th} I_C}{1 + D_2 V_A R_{th} I_C} \quad (5.7)$$

where $D\beta$ is the temperature coefficient of β. Typically, $D\beta V_A R_{th} I_C << D_2 V_A R_{th} I_C$ as $D\beta << D_2$. As can be seen from (5.7), the $(1 - \beta_o r_o g_\mu)$ term is always greater than zero.

From (5.6), for R_E and $R_B << r_\pi$, the difference between R_{out} and the output resistance predicted by electrical-only model, R_{outE}, is controlled by the degradation in r_o. In this case, measured output resistance can be significantly lower than ordinary SPICE predicts depending on the size of $D_2 R_{th} I_C V_A$. With increasing R_B and R_E, the error tends to be reduced. For R_B or $R_E >> r_\pi$, R_{out} stays about the same as R_{outE}; R_{out} drops only due to the variation in β with temperature which can be considered as a second-order thermal effect. For example, for the limiting case of $R_B \rightarrow 0$ and $R_E >> r_\pi$ (i.e., near-constant emitter current), (5.6) simplifies to

$$R_{out} \cong r_o \frac{1 + \beta_o}{1 - \beta_o r_o g_\mu} \cong \frac{R_{outE}}{1 + D\beta R_{th} I_C V_A} \quad (5.8)$$

where $R_{outE} = r_{oE}(1+\beta_o)$. As discussed above, typically, the $D\beta R_{th}I_C V_A$ term is substantially smaller than 1. In (5.8), the degradation in r_o is canceled out to first order by the $(1 - \beta_o r_o g_\mu)$ term.

The input resistance R_{in} of the same circuit is given by

$$R_{in} = r_\pi + R_E//R_C + \left[\frac{R_E + \beta_o(R_E + R_C)}{r_o + (R_E + R_C)(1 - \beta_o g_\mu r_o)}\right]\left(\frac{r_o R_E}{R_E + R_C} + r_\pi r_o g_\mu\right). \tag{5.9}$$

Note that this expression can be obtained from (5.6) by inspection, by exploiting the symmetry of the equivalent circuit. Notable in (5.9) is the strong dependence of R_{in} on the loading at the output port. For example, for the limiting case of $R_E \rightarrow 0$, (5.9) simplifies to

$$R_{in} \cong \frac{r_\pi}{1 - \beta_o g_\mu (r_o//R_C)}. \tag{5.10}$$

For $R_C << r_o$, (5.10) simplifies to $R_{in} \cong r_\pi$. Thermal effects on r_π are negligible. However, in the other limit $R_C >> r_o$, the denominator of (5.10) can be substantially smaller than 1, so that the input resistance can be many times greater than what the electrical-only model predicts. This effect can be explained as follows: an increment in the input voltage tends to increase the collector current; this incremental current ΔI_C reduces the collector-emitter bias. For large

enough R_C (i.e., $R_C >> r_o$), the power dissipation can be reduced even though the collector current has been increased. In that case, the base current increases less than that predicted by the electrical-only model, as the emitter-junction temperature actually drops, increasing the effective input resistance.

These thermally induced variations in input and output resistances lead to errors in other circuit responses such as voltage gain. The lower output resistance can significantly reduce the low-frequency voltage gain. In some cases, the increase in input impedance can partially cancel the effect of decreasing output impedance. In such a case, the net effect on gain can be small even when both impedances are strongly in error. Figure 5.16 shows an equivalent circuit of such case. In Fig. 5.16, R_S is the output resistance of previous stage, R_{in} the input resistance, g_m the transconductance, R_{out} the output resistance, and R_L the load resistance of amplifying stage. Under the conditions that $R_S >> R_{in}$ and $R_L >> R_{out}$ (that is, the gain stage is driven by a high impedance source and has a high impedance load), the voltage gain of this circuit is given by

$$A_v = \frac{g_m}{R_S} R_{in} R_{out}. \qquad (5.11)$$

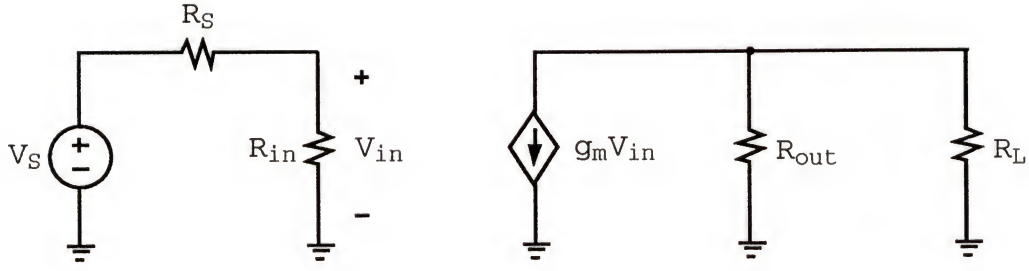


Fig. 5.16. Equivalent circuit diagram of a typical gain stage.

In (5.11), the higher R_{in} and the lower R_{out} are multiplied together so that the net value stays about the same as what the electrical-only model predicts. This is an important case as such situation happens in the gain stage of typical two-stage op-amps (e.g., 741). Even though there is little effect on the DC gain, errors in the input and output impedance can lead to significant errors in the frequency response of the circuit, as will be discussed later.

The small-signal aspects of thermal effects so far are demonstrated in following example. Figure 5.17 shows a common-emitter amplifier with a PNP active load. By inspection, the output resistance of this circuit is determined by the parallel combination of the output resistances of transistor Q_1 and Q_2. From the discussions above, the gain of this circuit can be considerably less than what electrical-only model predicts. Figure 5.18 shows the DC transfer characteristic of this circuit based on modified and ordinary SPICE simulations for $I_{C1} \cong 1$ mA. The figure shows that thermal feedback not only reduces the DC voltage gain but makes it a more strongly nonlinear function of the input swing. This is caused by the strong V_{CE} dependence of the current-mirror output I_{C2}. From the modified SPICE simulation, I_{C2} varies about 50% over the output swing.

To confirm the thermal effect on the input resistance, consider the circuit of Fig. 5.17 with the simple current

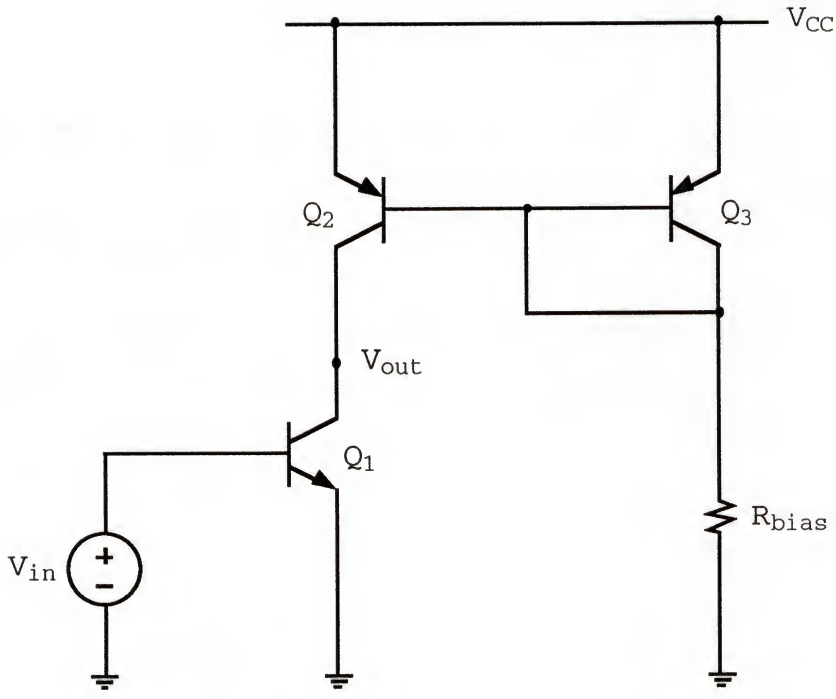


Fig. 5.17. Circuit diagram of a simple common-emitter with active-load.

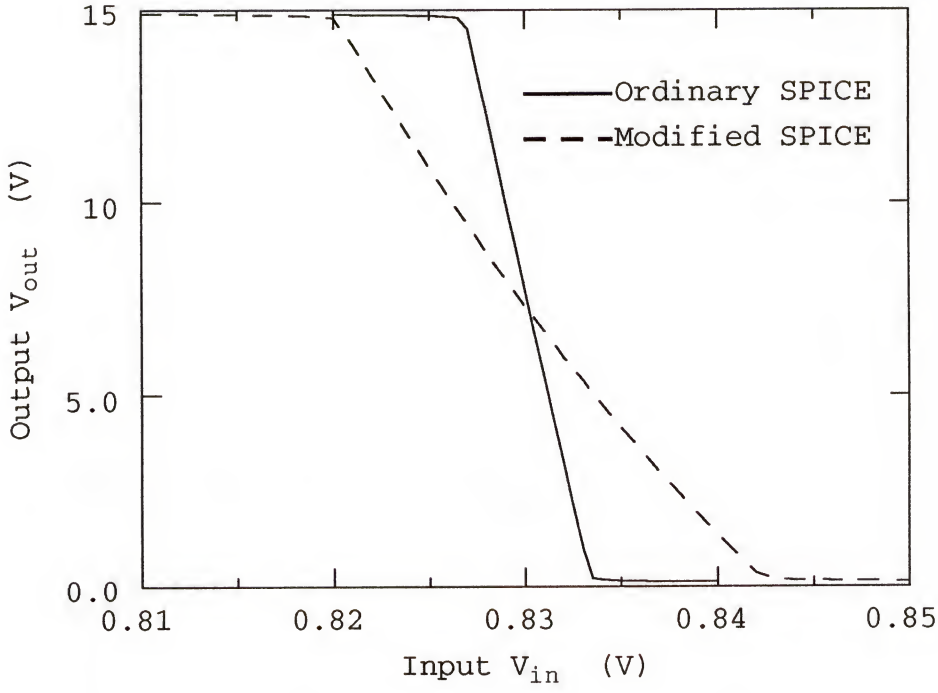


Fig. 5.18. Transfer characteristic of a simple common-emitter circuit.

mirror (Q_2 - Q_3) replaced with a Wilson current source (Fig. 5.19). In this case the output resistance of the circuit is given by (5.10). Modified SPICE predicted an input resistance of 1.5×10^4 Ω while ordinary SPICE gives 6.1×10^3 Ω, a prediction error of 146 %.

Figures 5.20a and b show the gain and phase responses of the circuit in Fig. 5.17. In both gain and phase, the results of simulations with and without thermal effects show considerable discrepancies. The low-frequency voltage gain predicted by modified SPICE is 1.49×10^3 while ordinary SPICE predicts 4.53×10^3, a prediction error of over 200 %.

Note that, in Fig. 5.20a, the voltage gain from modified SPICE varies at frequencies where ordinary SPICE predicts a constant value, especially near 1×10^6 Hz. As can be seen in (5.3), the frequency dependence of Z_{th} in y-parameters can cause the input impedance, output impedance, and the gain of a circuit to vary with frequency at frequencies where ordinary SPICE predicts a constant value. The frequency response of Z_{th} depends on the device structure [Fo91b]. The narrower the base-width and the emitter area, the higher the frequency over which thermal feedback will present. From Chapter 3, in a uniform medium, the time at which the impulse response at a point a distance r away reaches maximum is given as $t_{pk} = r^2/6K$. At t_{pk} the temperature increases to roughly 10% of its final value. Therefore, in principle, the

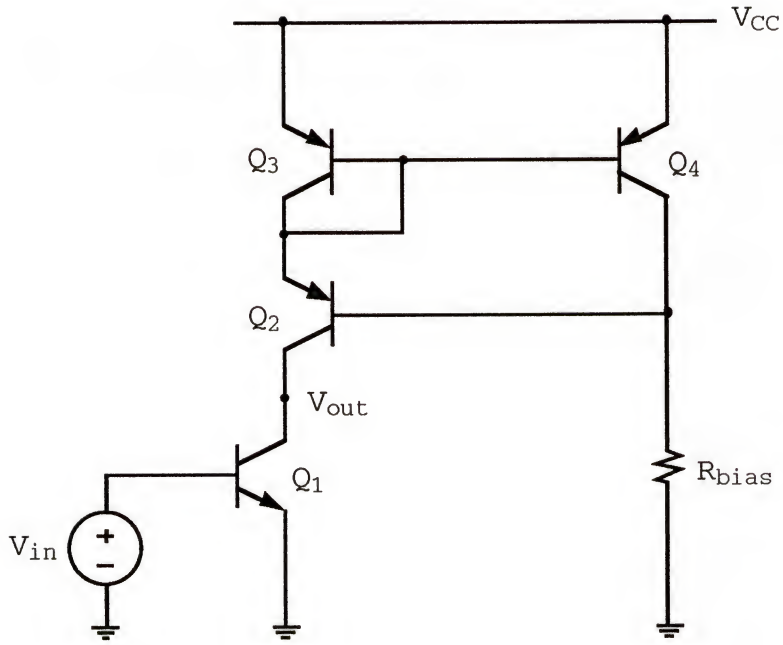


Fig. 5.19 Circuit diagram of a simple common-emitter with a Wilson source as active-load.

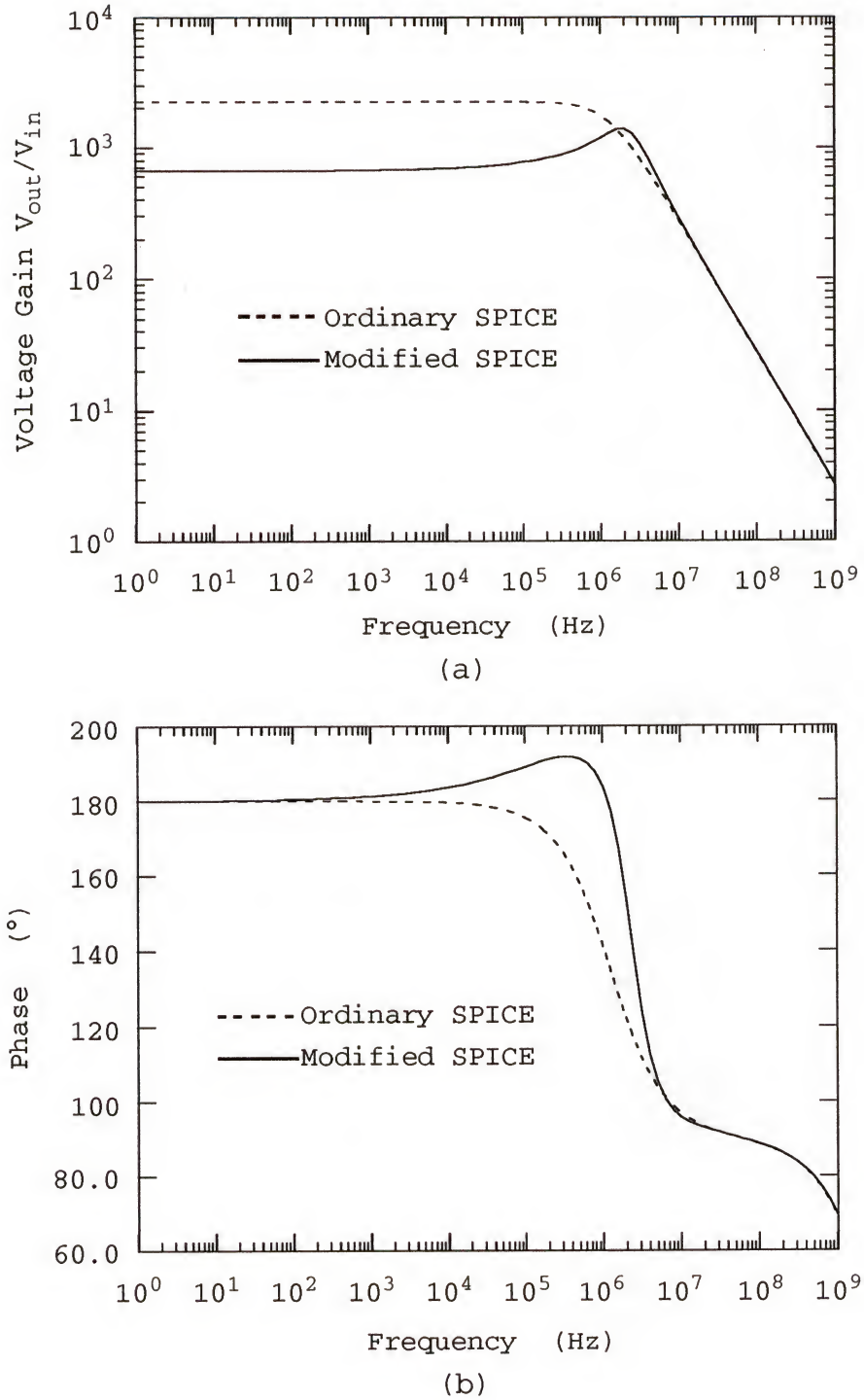


Fig. 5.20. Frequency response of a simple common-emitter circuit with active-load. (a) Gain. (b) Phase.

infinitesimal heat source located in collector-junction SCR will raise the temperature of the closest emitter junction to 10% of its final value within $W_B^2/6K$ seconds (W_B is metallurgical base-width), while for the farthest emitter junction, it takes $[(W_B+H)^2+W^2+L^2]/6K$ seconds. Typically, base-width is much smaller than emitter-width and length so that the emitter area is the dominating factor to determine t_{pk} of device, or the frequency characteristic of Z_{th}. From (5.1b) and (5.1c), $r_{eff} = 2\sqrt{WL}f_1f_2$.

In Fig. 5.20b, the hump in the phase response of the modified SPICE simulation is due to the negative phase of Z_{th} in Y_{22} as seen in Fig. 5.5b.

The stability of a circuit can be affected by the frequency dependence of the thermal impedance discussed above. At low frequencies, the stability of a circuit can also be affected more significantly by the thermally induced modulations in input and output resistances. The effect of the modulation in input and output resistance to the pole frequencies of a circuit will be considered here. Similar arguments can be applied to zeros also. Assume that the pole frequencies are low enough that (5.4) is valid and the collector current is high enough to cause significant thermal effects. Depending on the circuit configuration, four cases are possible. If pole frequency f_p is controlled by an output resistance, the smaller R_{out} due to the thermal effects can

increase f_p compared to the electrical-only model prediction. If this f_p was the dominant pole frequency, the circuit will have smaller phase margin than ordinary SPICE predicts, assuming the rest of the poles are not affected by thermal feedback or by the variation in this pole. The opposite can be true if this f_p was the frequency of a nondominant pole. For the case where the pole frequency f_p is controlled by an input resistance, which can be raised by self-heating, the thermal feedback moves the pole to lower frequency. Again, depending on whether the pole was dominant or not, this variation in f_p can increase or decrease the phase margin of the circuit. Table 1 summarizes the resulting effects for the four possible cases.

A change in pole frequency by less than an order of magnitude has little effect on the phase margin of the circuit if the rest of the poles were initially separated by more than two decades. At frequencies more than a decade from the pole frequency, phase effects are small. For Cases (a) and (d) of Table 1, if the pole influenced by thermal feedback was initially separated by less than two decades from some other pole, thermal feedback can degrade the phase margin. In Cases (b) and (c), the circuit will be over-compensated. Considering the chip area that will be taken up by the compensation capacitors, accurate prediction of

Table 1: List of various conditions under which thermal feedback can affect the stability of a circuit resulting from the discrepancies in the pole frequency f_p.

Effect on f_p	If pole constitutes	Resulting effects
Smaller R_{out} increases f_p	(a) Dominant pole	Reduces phase margin
	(b) Nondominant pole	Increases phase margin
Larger R_{in} reduces f_p	(c) Dominant pole	Increases phase margin
	(d) Nondominant pole	Reduces phase margin

frequency response is required to optimally size the compensating capacitor.

Note, as discussed in the DC aspects, correcting the forward Early voltage V_A cannot solve the problems induced by thermal effects. If V_A is extracted at low current, simulation results are accurate at low current for all frequencies and at high frequencies for all currents. However, simulation results will be wrong at low frequencies with moderate to high currents. On the other hand, if V_A is extracted at high currents, simulation results will be accurate at high currents for low frequencies but wrong at low currents for all frequencies and at high frequencies for all currents. Either way, designers cannot avoid errors unless thermal effects are included properly in the simulator. Table 2 summarizes the above discussions.

A possible mistake in circuit compensation by not understanding the aspect summarized in Table 2b is illustrated in the following example. Figure 5.21 shows a two-stage amplifier in which the second stage will generate a pole at

$$|p_1| = \frac{1}{R_p C_p} \tag{5.12}$$

where R_p and C_p are the capacitance and resistance of the second stage, respectively. For $R_C >> r_o$, based on the BJT

Table 2. Summary of the relations between ordinary SPICE simulations error and operating conditions depending on whether the Early voltage is extracted at high currents or low currents. Extracting V_A at low currents represents V_A considering only the electrical effects, high frequencies represents high enough frequencies to avoid thermal effects, and low currents represents negligible power dissipation.

(a) V_A extracted at low I_C

Simulation results	Low currents	High currents
Low frequencies	Accurate	Wrong
High frequencies	Accurate	Accurate

(b) V_A extracted at high I_C

Simulation results	Low currents	High currents
Low frequencies	Wrong	Accurate
High frequencies	Wrong	Wrong

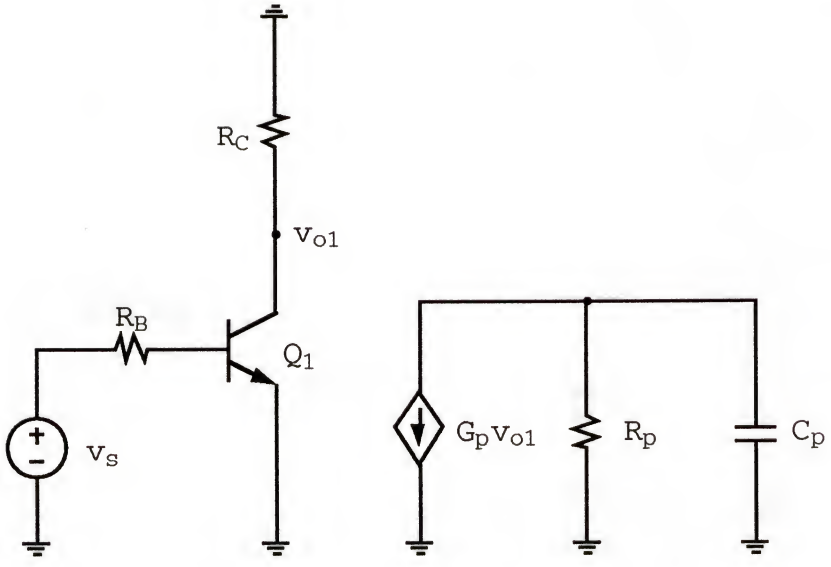


Fig. 5.21. Small-signal equivalent circuit diagram of a two stage amplifier.

equivalent circuit shown in Fig. 5.15b, the first stage has a pole near

$$|p_2| = \frac{1}{c_\pi(R_B//r_\pi) + c_{cs}\,r_o + c_\mu\left[\dfrac{r_o + g_m r_o(R_B//r_\pi)}{1 - g_m r_o g_\mu(R_B//r_\pi)}\right]} \qquad (5.13)$$

where c_π, c_μ, and c_{cs} are the parasitic capacitances of transistor Q_1. Let's assume $|p_1| < |p_2|$, and $|p_2|$ is high enough that thermal effects are negligible at that frequency, and assume this circuit need to be compensated for unity gain with a phase-margin of 45°. One could add a capacitance C_p' in parallel with C_p to lower the dominant pole frequency. C_p' must satisfy the relation

$$C_p' + C_p = \frac{A_{vo}}{R_p\,|p_2|} = \frac{G_{m1}R_{o1}G_p}{|p_2|} \qquad (5.14)$$

where A_{vo} is the low-frequency voltage gain of the circuit, G_{m1} the transconductance of first stage, R_{o1} the output resistance of first stage. $|p_2|$ and R_{o1} must be known to estimate the size of C_p' and designers typically use simulation. From (5.13), $|p_2|$ is a function of r_o, which is directly proportional to V_A, and $R_{o1} = r_o$. Suppose V_A of Q_1 was extracted at the actual operating current so that the low-frequency gain is predicted correctly. Figure 5.22a and b compare the frequency response of this circuit based on both ordinary and modified SPICE simulation. In Fig. 5.22, V_A = 56

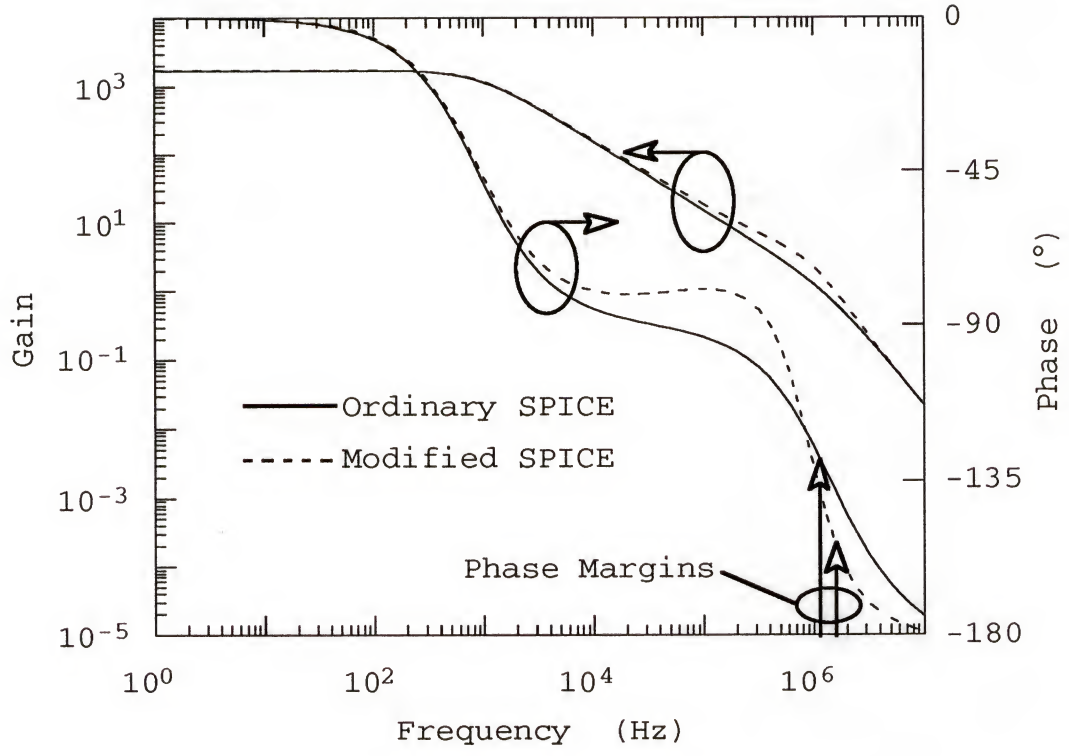


Fig. 5.22. Gain and phase versus frequency for a two stage amplifier.

V, which was extracted at I_C = 1 mA, was used for the ordinary SPICE simulation and C_p' was adjusted to give 45° phase-margin at unity-gain based on (5.13) and (5.14). The same C_p' was used for modified SPICE simulation, but V_A was left to be 150 V, which is the Early voltage considering electrical effects only, as in real circuit. As can be seen in Fig. 5.22a, ordinary SPICE gives the same voltage gain as what modified SPICE does at low frequencies. However, with increasing frequency, the actual voltage gain rises above the prediction of ordinary SPICE as thermal effects diminish. At frequencies high enough to neglect thermal effects, output resistance r_o is controlled by the electrical effects only. Therefore, using V_A = 56 V leads to underestimation of C_p'. This circuit designed to have 45° ended up getting only about 25° phase margin.

5.4. Summary

Depending on the bias conditions, thermal feedback increases the terminal currents (I_E, I_B, and I_C) or reduces the base-emitter voltage of BJTs. These DC thermal effects can lead to significant errors in analog circuits as they make the terminal currents more strongly dependant on collector-emitter voltage and to be a less-exact exponential function of base-emitter voltage. Many high-accuracy analog

circuits, especially translinear circuits, are susceptible to these DC thermal effects.

In low-frequency analysis, thermal feedback significantly increases the conductances g_{12} and g_{22}. Increase in g_{22} leads to significant reduction in output resistance of BJT circuit for the case transistor is driven by a low impedance source, while the increase in g_{12} increases the input resistance of the BJT circuit for $R_C >> r_o$. Thermal effects on input and output resistances can lead to errors in gain. In addition, the frequency dependence of thermal impedance gives an additional frequency dependence in input impedance, output impedances, and gain, which is not predicted by the existing small-signal model. Discrepancies in input and output impedances, both at low and high frequency, can lead to errors in stability prediction of the integrated circuit.

Reducing the maximum collector-emitter voltage of the transistors by employing smaller supply voltages or cascoding can reduce thermal effects on the parameters that are affected by the power dissipation. However, Y_{12} and Y_{22} can be significantly affected independent of power dissipation. Thermal effects on terminal currents and output resistance tend to be less significant with increasing base or emitter resistance as the negative electrical feedback across these

resistors partially offsets the positive thermal feedback effects.

Note that as discussed in previous chapters, errors such as those in the foregoing examples will be observed at much lower currents in dielectrically isolated (DI), SOI, or trench-isolated transistors. For example, the DI BJTs had about four times higher thermal resistances than the silicon junction isolated transistors [Fo91b]. Therefore, for DI devices, the errors presented in Section 5.3 can be seen at I_C less than 250 μA. The predictive thermal impedance model presented in Chapter 3 is not directly valid for DI, SOI, and trench isolated devices. This model was developed for homogeneous materials. However, since the DC simulations solely depend on the magnitude of thermal resistance, modified SPICE can be used for the DC simulations of any structure as long as the thermal resistance is available.

CHAPTER 6
SUMMARY AND SUGGESTIONS FOR FUTURE WORK

In this dissertation, predictive modeling of high-current output resistance, and thermal effects on modeling, characterization, and circuit design, which are increasingly significant with device scaling and isolation technologies using dielectric material, have been discussed. A study of the effects of carrier-velocity saturation on the output resistance characteristic of BJTs was done. Physical analysis demonstrates that BJT output resistance can be raised significantly by carrier velocity-saturation especially at high currents. Analytical expressions for normalized BJT output resistance helps to understand the effects qualitatively. On the thermal part, this work demonstrated the increasing significance of local-heating effects with device scaling and identified the roles of thermal spreading impedance and chip/package impedance on BJT thermal behavior depending on the power dissipation. The physical analysis of BJT behavior corrected for temperature variation demonstrates the significance of thermal effects on BJT small-signal parameters, which justifies the use of frequency-domain thermal model. The thermal impedance model is simple, yet

predictive, and can be used for compact modeling in SPICE to improve the accuracy of BJT circuit simulation. The fundamental limits with standard thermal parameter extraction methods are pointed out, and new methods are introduced which are compatible for any type of device, including advanced BJT structures. From the study of circuit implications of self-heating, this work provided general insight into thermal effects on integrated circuit design, identified the aspects of thermal effects that can lead to significant errors in a certain circuit configuration, and provided design principles to reduce thermal effects in integrated circuit design.

For future work, the first suggestion goes to the exploration of SPICE BJT model for temperature dependences of the SPICE model parameters. Increase in device temperature tends to be really significant at currents where the transistor goes into high level injection, and the device performance strongly depends on the temperature dependences of the terminal currents and voltages. This work relied on the temperature models used in SPICE to predict the thermal effects on high currents. However, the accuracy of these models has not been fully explored. Considering the empiricism prevailing in the SPICE BJT model, it is most likely that more physical models for temperature coefficients are required for better prediction of high-current thermal behavior.

Transient thermal effects have not been discussed much in this work. One of the reasons was that, based on Denison and Walter's work [De89], local-heating was expected to have little effect on digital circuit performance. Another reason was that a transient thermal circuit simulator was not available. Thermal effects can significantly degrade the settling time of a typical output buffer with large voltage pulse applied at the input, or a switching stage where the voltages are switched by activating/deactivating current sources. Therefore, transient thermal effects need to be explored on analog circuit stand point, and in order to do that, it is important to have the simulator available.

Thermal effects are far more significant in SOI, dielectrically isolated, and trench isolated transistors. Devices built with these processes tend to be used in many high performance analog circuits because of the advantages provided by them (e.g., lower collector-junction capacitance and lower isolation leakage currents at high temperature). Measurements indicate that, for dielectrically isolated devices, thermal effects are about four times more significant but disappear at lower frequencies than same-dimension junction-isolated devices. Modeling thermal impedance for these devices will be important and very useful for both analog and digital circuit simulations.

REFERENCES

[Ar74] R. P. Arnold and P. L. Zoroglu, "A quantitative study of emitter ballasting," IEEE Trans. Electron Devices, ED-21, pp. 385, 1974.

[Br74] A. P. Brokaw, "A simple three-terminal bandgap reference," IEEE J. Solid-State Circuits, vol. SC-9, pp. 388-393, Dec. 1974.

[Ca62] H. S. Carslaw and J. C. Jaeger, Conduction of heat in solids, 2nd ed. Oxford: Oxford University Press, 1962.

[Da88] W. F. Davis and M. L. Lidke, "The effect of thermal feedback within the bipolar transistor on Gummel-Poon model accuracy," Motorola Internal Report, May 27, 1988.

[De89] R. T. Denison and K. M. Walter, "Local thermal effects in high performance bipolar device/circuits," Proc. IEEE Bipolar Circuits and Technology Meeting, pp. 164-167, 1989.

[Fo91a] R. M. Fox and S.-G. Lee, "Predictive modeling of thermal effects in BJTs," Proc. IEEE 1991 Bipolar Circuits and Technology Meeting, pp. 89-92, Sept. 1991.

[Fo91b] R. M. Fox and S.-G. Lee, "Thermal parameter extraction for bipolar circuit modeling," Electronics Letters, vol. 27, no. 19, pp. 1719-1720, Sept. 1991.

[Fo91c] R. M. Fox and S.-G. Lee, "Scalable small-signal model for BJT self-heating," IEEE Electron Device Lett., vol. 12, no. 12, pp. 649-651, Dec. 1991.

[Ge78] I. E. Getreu, Modeling the bipolar transistor. New York: Elsevier, 1978, p 226.

[Ho58] R. Holm, Electrical contacts handbook. Berlin: Springer-Verlag, 1958.

[Ho74] P. L. Hower and P. K. Govil, "Comparison of one- and two-dimensional models of transistor thermal instability," IEEE Trans. Electron Devices, vol. ED-21, pp. 617, 1974.

[Je91] H. Jeong, "MMSPICE: A physical alternative to Gummel-Poon/SPICE and a semi-numerical mixed-mode simulator for advanced bipolar technology CAD," Ph. D. dissertation, University of Florida, Gainesville, FL, Dec. 1989.

[Jeo89] H. Jeong and J. G. Fossum, "A charge-based large-signal bipolar transistor model for device and circuit simulation," IEEE Trans. Electron Devices, vol. ED-36, no. 1, pp. 124-131, Jan. 1989.

[Jo70] R.C. Joy and E. S. Schlig, "Thermal properties of very fast transistors," IEEE Trans. Electron Devices, vol. ED-17, no. 8, pp. 586-594, August 1970.

[Ki62] C. T. Kirk, Jr., "A theory of transistor cutoff frequency (f_t) fall-off at high current densities," IRE Trans. Electron Devices 9, pp. 164, 1962.

[Ku73] K. E. Kujik, "A precision reference voltage source," IEEE J. Solid-State Circuits, vol. SC-8, pp. 222-226, June 1973.

[La81] M. Latif and P. R. Bryant, "Multiple equilibrium points and their significance in the second breakdown of bipolar transistors," IEEE J. Solid-State Circuits, vol. SC-16, no. 1, February 1981.

[Le92] S.-G. Lee and R. M. Fox, "The effects of carrier-velocity saturation on high-current BJT output resistance," IEEE Trans. Electron Devices, vol. ED-39, no. 3, pp. 629-633, March 1992.

[Li90] J. J. Liou, "Comments on 'Early voltage in very-narrow-base transistors," IEEE Electron Device Lett., vol. EDL-11, no.5, p.236, May 1990.

[Me77] G. Meijer, "The current dependency of the output conductance of voltage-driven bipolar transistors,"

IEEE J. Solid-State Circuits, vol. SC-12, no. 4, pp. 428-429, August 1977.

[Mu64] O. Muller, "Internal thermal feedback in four-poles especially in transistors," Proc. IEEE, vol. 52, pp. 924-930, Aug. 1964.

[Ot76] F. Ottinger, D. L. Blackburn, and S. Rubin, "Thermal characterization of power transistors," IEEE Trans. Electron Devices, vol. ED-23, no. 8, pp. 831-838, August 1976.

[Pi85] M. R. Pinto, C. S. Rafferty, H. R. Yeager, and R. W. Dutton, "PISCES-IIB user's manual and supplement report," Stanford Electronics Lab. Tech. Rep., Stanford University, Stanford, CA, 1985.

[Po89] Pong-Fei Lu, and Tze-Chiang Chen, "Collector-base junction avalanche effects in advanced double-poly self-aligned bipolar transistors," IEEE Trans. on Electron Devices, vol. 36, no. 6, pp. 1182-1188, June 1989.

[Ps89] PSpice, MicroSim Corporation, 20 Fairbanks, Irvine, California 92708, July 1989.

[Re92] Michael Reisch, "Self-heating in BJT circuit parameter extraction," Solid-State Electronics, vol. 35, no. 5, pp. 677-679, 1992.

[Ro90] D. J. Roulston, "Early voltage in very-narrow-base bipolar transistors," IEEE Electron Device Lett., vol. EDL-11, pp. 88-89, 1990.

[Sc67] H. A. Schafft, "Second breakdown-A comprehensive review," Proc. IEEE, vol. 55, pp. 1272-1288, August 1967.

[Sl76] J. W. Slotboom and H. C. de Graaff, "Measurements of bandgap narrowing in Si bipolar transistors," Solid-State Electron., vol. 19, pp. 857-862, October 1976.

[Sp58] J. J. Sparks, "Voltage feedback and thermal resistance in junction transistors," Proc. IRE, vol. 46, pp. 1305-1306, June 1958.

[St59] P. R. Strickland, "The thermal equivalent circuit of a transistor," IBM J. Res. & Dev., vol. 3, pp.35-45, January 1959.

[Sz81] S. M. Sze, Physics of semiconductor devices. New York: John Wiley & Sons, Inc., 1981.

[Vo89] R. Vogelsong and C. Brzezinski, "Simulation of thermal effects in electrical systems," APEC '89, Baltimore, MD, March 1989.

[Ye91] F. Q. Ye, "A BJT model with self heating for WATAND computer simulation," M.S. Thesis, Youngstown State University, Youngstown, OH, June 1990; also described in IEEE Circuits and Devices, vol. 7, no. 3, pp. 7-9, May 1991.

[Zd87] P. J. Zdebel, R. J. Balda, B.-Y. Hwang, V. d.t. Torre, and A. Wagner, "MOSAIC III - A high performance bipolar technology with advanced self-aligned devices," IEEE Bipolar Circuits and Technology Meeting, pp. 172-175, 1987.

[Zw92] D. T. Zwdinger, S.-G. Lee, and R. M. Fox, "Compact modeling of BJT self-heating in circuit simulation," Proceedings of the NUPAD IV, Seattle, June 1992.

BIOGRAPHICAL SKETCH

Sang-Gug Lee was born in Kyung-Nam, Korea, on August 3, 1958. He received the B.S. degree in electronic engineering from KyungPook National University, TaeGu, Korea, in 1981. From 1983 to 1986 he was with the Naval Academy, Jin-Hae, Korea, where he was an instructor in the department of electronic engineering. He received the M.E. degree in electrical engineering from University of Florida, Gainesville, FL, in 1989. Since 1989, he has been working toward the Ph.D degree at the University of Florida, Gainesville, specializing in BJT modeling, characterization, and circuit design.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox, Chairman
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jerry G. Fossum
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Dorothea E. Burk
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Eisenstadt
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kevin S. Jones
Associate Professor of Materials Science and Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1992

Hubert A. Bauer
for Winfred M. Phillips
Dean, College of Engineering

Madelyn M. Lockhart
Dean, Graduate School